G.I. EPIFANOV

ÉLECTRONIQUE DES SOLIDES

G.I. EPIFANOV
Y. MOMA

ÉLECTRONIQUE DES SOLIDES

AVANT-PROPOS

Le présent manuel est écrit sur la base d'un cours de conférences que les auteurs ont lues pendant plusieurs années à l'Institut des constructions électroniques de Moscou aux étudiants qui se spécialisent dans le domaine de la construction des machines électroniques, des matériaux utilisés dans la technique électronique, de la construction et de la technologie des appareils radio-électroniques et de calcul électronique, des appareils d'automatisation et de calcul automatique, de la technologie des appareils à semiconducteurs et des circuits intégrés.

Le présent cours est un cours *d'introduction* qui est professé aux premiers semestres et suit directement le cours de physique générale dans lequel les étudiants reçoivent des connaissances élémentaires sur la mécanique quantique et la statistique physique qui sont nécessaires pour l'assimilation de ce cours.

Le cours se propose d'atteindre un double but. Premièrement, il décrit d'une façon systématique, bien que succincte, les phénomènes physiques qui sont nécessaires pour la compréhension des principes de fonctionnement des principaux types d'appareils à solides: il donne des idées générales sur la structure des cristaux parfaits et réels, les propriétés thermiques des solides, la théorie des bandes, la statistique des porteurs de charges en équilibre et en non-équilibre dans les métaux et les semiconducteurs, la conductibilité électrique des métaux et des semiconducteurs, les phénomènes galvanomagnétiques, thermomagnétiques et thermoélectriques, les phénomènes plasmiques et optiques dans les semiconducteurs, les propriétés magnétiques des solides.

Deuxièmement, pour satisfaire aux besoins des disciplines spéciales ultérieures, le cours contient une étude sommaire de la physique des phénomènes de contact, une théorie élémentaire de la jonction *p-n* et les principes physiques, découlant de cette étude, de fonctionnement des principaux types d'appareils à semiconducteurs: des diodes (redresseuses, impulsionnelles, à haute fréquence, à effet tunnel, à avalanche et autres), des appareils photoélectriques, des transistors bipolaires et à effet de champ, des appareils à caractéristique courant-tension en S, ainsi que les idées les plus générales sur les circuits intégrés et l'utilisation des appareils à semiconducteurs comme éléments de ces circuits.

Dans l'exposé, une place primordiale est accordée à la mise en évidence de la physique des phénomènes et des processus, souvent sans donner de justifications et de démonstrations mathématiques rigoureuses. Comme l'a montré la pratique de plusieurs années de lecture d'un tel cours complété par la résolution des problèmes et l'exécution de calculs personnels qui fait connaître les principales idées physiques et les principes de l'électronique des solides, ce cours est utile et a fait ses preuves dans les écoles d'enseignement supérieur soviétiques.

Les auteurs sont reconnaissants au rectorat de l'Institut des constructions électroniques de Moscou qui a recommandé d'éditer le manuscrit en russe ; aux collaborateurs de la chaire de bases physiques de la technique électronique pour une discussion utile du contenu du manuscrit et de la méthode de son exposé ; aux critiques scientifiques de l'édition russe de l'ouvrage — à la chaire de physique du solide de l'Université d'Etat de la Biélorussie et au docteur ès sciences physiques et mathématiques, le professeur T. Chermergor qui a bien voulu relire le manuscrit et faire des remarques critiques précieuses améliorant le contenu et le style de l'exposé de certaines parties du livre ; à Z. Epifanova et A. Katounina pour leur travail non facile relatif à la présentation du manuscrit ; aux éditions « Vyschaïa chkola » pour l'édition du livre en russe ; aux éditions « Mir » qui ont bien voulu se charger de l'édition du livre en français.

CHAPITRE PREMIER

STRUCTURE DES CRISTAUX

La majorité absolue des appareils à solides sont construits à base de corps monocristallins aussi purs que possible et, en premier lieu, de monocristaux semiconducteurs localement dopés en impuretés déterminées. Aussi, pour comprendre la physique de fonctionnement de tels appareils est-il nécessaire d'avoir des idées précises sur la structure des cristaux, les méthodes de sa description et de son étude expérimentale, ainsi que sur les forces de liaison qui déterminent cette structure. C'est à une étude sommaire de ces questions qu'est consacré le chapitre actuel.

§ 1.1. Réseau cristallin

Lorsque les atomes et les molécules se rapprochent, il se crée entre eux des forces d'interaction. Leur caractère général est montré sur la fig. 1.1, *a*: à des distances relativement grandes, il y a apparition des forces d'attraction F_{att} qui croissent lorsque la distance r entre les particules diminue (courbe *1*); à de petites distances apparaissent des forces de répulsion $F_{rép}$ qui croissent, lorsque r diminue, beaucoup plus rapidement que F_{att} (courbe *2*). A une distance $r = r_0$, les forces de répulsion équilibrent les forces d'attraction et la force résultante F (courbe *3*) s'annule tandis que l'énergie d'interaction U des particules prend sa valeur minimale égale à $-U_l$ (fig. 1.1, *b*). C'est pourquoi l'état des particules séparées par la distance r_0 est un état d'équilibre stable; de ce fait, les particules doivent s'arranger sous l'action des forces d'interaction dans un ordre rigoureux, en constituant un corps de structure interne régulière — un *cristal*. Pour décrire une telle structure il est commode d'utiliser la notion de *réseau cristallin*. On distingue les *réseaux de translation de Bravais* et les *réseaux à base*.

Réseaux de Bravais. Au point de vue géométrique l'arrangement ordonné et périodique des particules dans le cristal peut être décrit à l'aide de l'*opération de translation*. La fig. 1.2, *a* montre un réseau obtenu par translation d'une particule le long de trois axes: le long de l'axe OX de segments a, $2a$, $3a$, ..., ma, le long de l'axe OY de segments b, $2b$, $3b$, ..., nb et le long de l'axe OZ de segments c, $2c$, $3c$, ..., pc, où m, n, p sont des nombres entiers. La position de toute particule dans un tel réseau est définie par le vecteur

$$\mathbf{r} = m\mathbf{a} + n\mathbf{b} + p\mathbf{c}. \tag{1.1}$$

Les vecteurs **a**, **b**, **c** sont les *vecteurs principaux de translation* et leurs modules sont les *périodes de translation*.

Le réseau construit par translation d'un nœud quelconque suivant les directions de translation est appelé réseau *de translation* ou réseau *de Bravais*. On appelle *maille élémentaire* le parallélépipède le plus petit construit sur les vecteurs **a**, **b**, **c** (fig. 1.2, *b*). Toutes les mailles élémentaires constituant le réseau ont une même forme et un même volume; à tous les sommets des mailles se situent des atomes ou des groupes d'atomes identiques. Aussi, tous les sommets des mailles sont-ils équivalents l'un à l'autre. On les appelle *nœuds* du réseau.

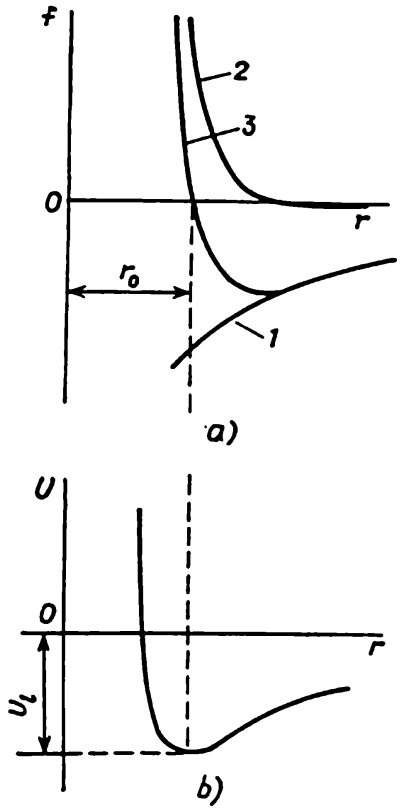

Fig. 1.1. Variation de la force (*a*) et de l'énergie *U* d'interaction (*b*) des particules en fonction de la distance *r* qui les sépare.

Pour caractériser une maille élémentaire il faut, dans le cas général, donner six quantités: trois arêtes de maille *a*, *b*, *c* et trois angles α, β, γ que ces arêtes font entre elles. Ces quantités sont dites *paramètres* du réseau. Souvent on prend pour unité de longueur dans les réseaux les segments *a*, *b*, *c*; on les appelle unités *axiales*.

Les mailles élémentaires contenant un seul nœud sont dites *simples* ou *primitives*. Leur choix n'est pas univoque. La figure 1.3 illustre à titre d'exemple les divers procédés de choix d'une telle maille. A chaque nœud du réseau correspond un même volume. Il est égal au volume de la maille élémentaire primitive et ne dépend pas du procédé de son choix.

Dans certains cas, pour obtenir une expression plus complète de la symétrie du réseau, les mailles élémentaires sont construites de telle sorte qu'elles contiennent des nœuds non seulement aux sommets mais aussi en d'autres points. De telles mailles sont dites *complexes*. Les plus répandues sont des mailles *centrées* (C), des mailles *à faces centrées* (FC) et des mailles *à base centrée* (BC). La figure 1.4 montre en pointillé les mailles à faces centrées (FC), centrée (C) et à base centrée (BC) et en traits pleins, les mailles primitives de ces réseaux. On voit que tous ces réseaux sont des réseaux de Bravais bien que pour la description de la symétrie du cristal il soit plus commode d'utiliser des mailles complexes et non pas simples.

On peut démontrer rigoureusement qu'il existe 14 et seulement 14 réseaux de Bravais différents qui se divisent suivant le degré de symétrie en 7 systèmes ou syngonies: *triclinique* dont la maille élémentaire est un parallélépipède oblique avec les côtés $a \neq b \neq c$ et les angles $\alpha \neq \beta \neq \gamma$; *monoclinique* qui est un parallélépipède

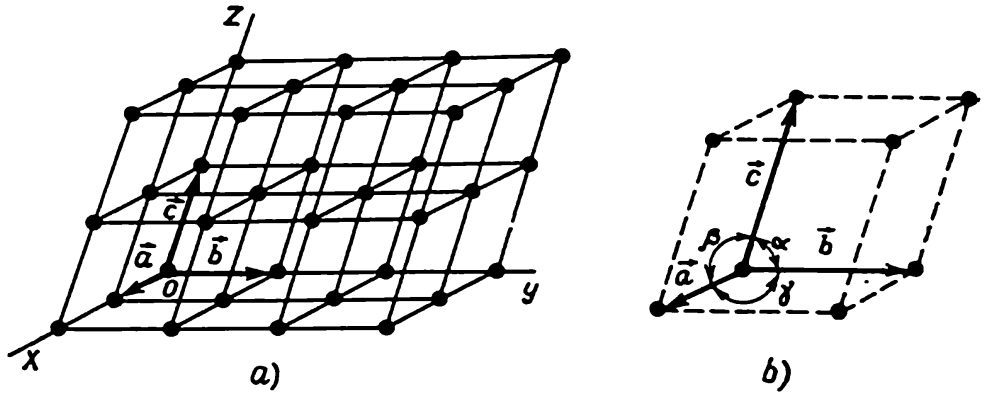

Fig. 1.2. Réseau de translation simple (*a*) et maille élémentaire de ce réseau (*b*): a, b, c — principaux vecteurs de translation; a, b, c, α, β, γ — paramètres de la maille élémentaire.

dont $a \neq b \neq c$, $\alpha = \gamma = 90°$, $\beta \neq 90°$; *orthorhombique* qui est un parallélépipède dont $a \neq b \neq c$, $\alpha = \beta = \gamma = 90°$; *tétragonal* (ou quadratique) qui est un parallélépipède dont $a = b \neq c$, $\alpha = \beta = \gamma = 90°$; *trigonal* qui est un parallélépipède dont $a = b = c$,

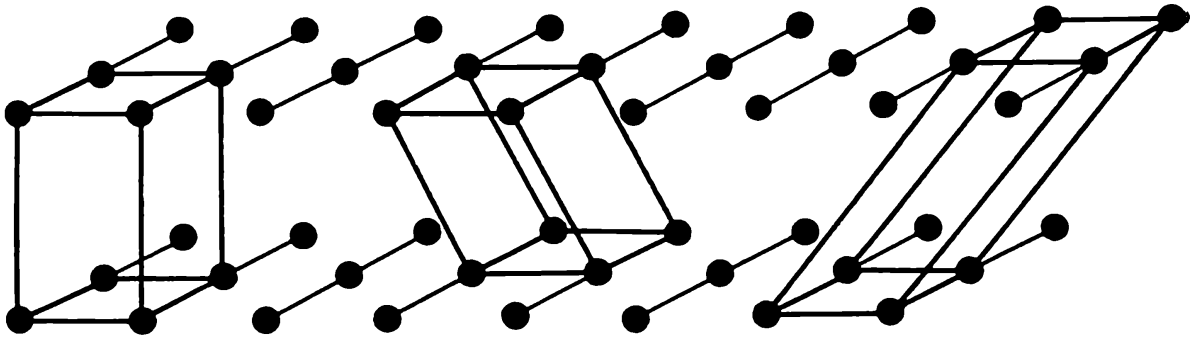

Fig. 1.3. Divers procédés de choix d'une maille élémentaire primitive.

$\alpha = \beta = 90°$, $\gamma \neq 90°$; *hexagonal* qui est un parallélépipède dont $a = b \neq c$, $\alpha = \beta = 90°$; $\gamma = 120°$; *cubique* qui est un parallélépipède dont $a = b = c$, $\alpha = \beta = \gamma = 90°$.

La figure 1.4 montre les mailles élémentaires de ces réseaux et leurs paramètres. Les vecteurs principaux sont figurés en traits pleins et les lignes auxiliaires en pointillé. Les mailles le moins symétriques sont celles du système triclinique et les mailles le plus symétriques sont celles du système cubique.

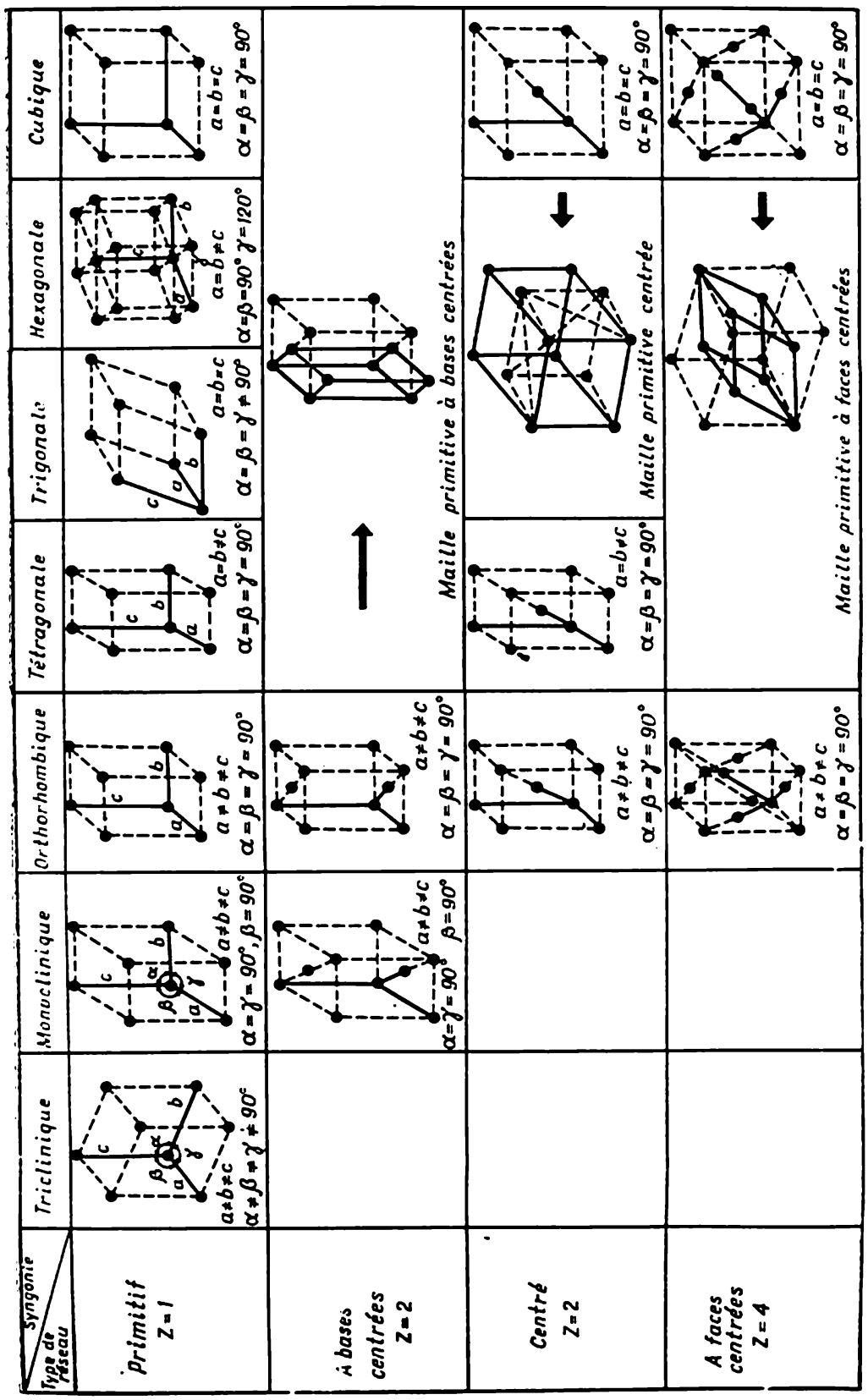

Fig. 1.4. Mailles élémentaires des quatorze réseaux spatiaux de Bravais divisés, suivant le degré de symétrie, en 7 systèmes ou syngonies et, suivant le type, en mailles simples ou primitives et mailles complexes ou multiples : à base centrée, centrées et à faces centrées.

Réseaux à base. Tout réseau ne peut pas être obtenu par la translation d'un seul nœud. La figure 1.5, *a* montre à titre d'exemple un réseau bidimensionnel de type général. Il est facile de voir qu'en construisant un tel réseau à l'aide des vecteurs de translation **a**, **b**, il est impossible d'obtenir une maille ne contenant qu'un seul nœud.

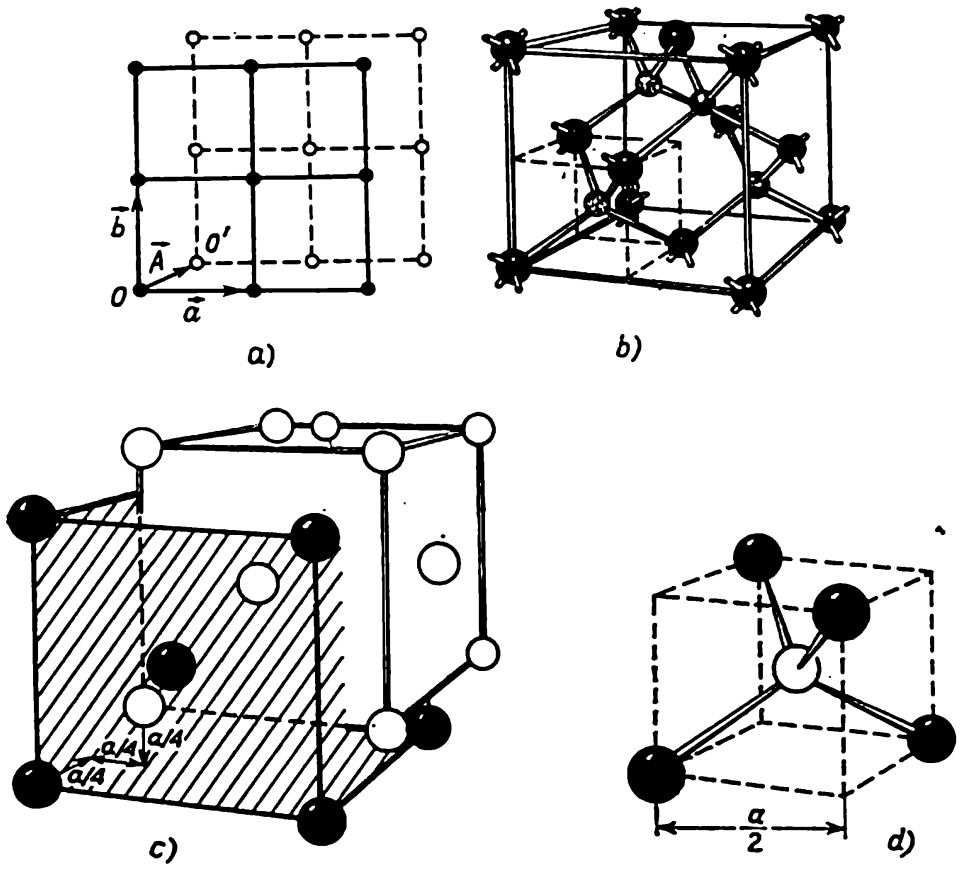

Fig. 1.5. Réseaux à base:

a — réseau plan à base (*OO'* — base du réseau); *b* — réseau tridimensionnel à base — réseau du type diamant à disposition tétraédrique des atomes (le tétraèdre est figuré en pointillé); *c* — un réseau du type diamant peut être constitué de deux réseaux cubiques à faces centrées emboîtés l'un dans l'autre et déplacés, suivant la diagonale spatiale, de 1/4 de la longueur de la diagonale; *d* — chaque atome de carbone est entouré par quatre voisins les plus proches situés aux sommets du tétraèdre.

Un tel réseau peut être représenté sous forme de deux réseaux de Bravais qui s'emboîtent l'un dans l'autre et dont chacun se détermine par les vecteurs de translation **a** et **b**. Le décalage des réseaux l'un par rapport à l'autre est décrit par un vecteur supplémentaire **A** qu'on appelle vecteur *de base*. Dans le cas général, le nombre de tels vecteurs peut être quelconque. Toutefois, on les prend d'ordinaire en nombre minimal nécessaire pour la construction du réseau.

Un réseau du type général est appelé réseau *à base*. Il peut être construit au moyen des mêmes translations qu'on utilise pour chacun des réseaux de Bravais constitutifs mais en effectuant la translation

non d'un nœud mais d'une base donnée par l'ensemble des vecteurs de base. C'est ainsi que le réseau montré sur la figure 1.5, *a* peut être obtenu par la translation de la base constituée par les nœuds *O* et *O'*. Comme exemple de réseau à base tridimensionnel, la figure 1.5, *b* montre le réseau de diamant. Il peut être formé par deux réseaux cubiques à faces centrées emboîtés l'un dans l'autre et décalés suivant la diagonale spatiale de 1/4 de longueur de la diagonale (fig. 1.5, *c*). Chaque atome de carbone est entouré par quatre voisins les plus proches situés aux sommets d'un tétraèdre montré en pointillé sur la figure 1.5, *b* et représenté sur la figure 1.5, *d*.

§ 1.2. Désignation des nœuds, des directions et des plans dans le cristal. Indices de Miller

Proposons-nous d'étudier sommairement les notations communément admises pour les nœuds, les directions et les plans dans le réseau cristallin.

Indices des nœuds. La position de tout nœud du réseau est déterminée par la donnée de trois coordonnées (fig. 1.6, *a*) : x, y, z. Pour un réseau simple ces coordonnées peuvent être exprimées comme suit :

$$x = ma, \quad y = nb, \quad z = pc, \tag{1.2}$$

où a, b, c sont les *paramètres* du réseau ; m, n, p des nombres entiers. Si l'on prend pour unité de longueur les paramètres du réseau, les coordonnées du nœud seront représentées par les nombres m, n, p. Ces nombres sont appelés *indices du nœud* et écrits sous la forme suivante : $[[mnp]]$. Pour des mailles à plusieurs nœuds les nombres m, n, p ne sont pas nécessairement entiers.

Indices de la direction. Pour la description d'une direction dans le cristal, on choisit une droite passant par l'origine des coordonnées. Sa direction se détermine univoquement par les indices $[[mnp]]$ du premier nœud par lequel elle passe (fig. 1.6, *a*). Aussi, les indices du nœud sont-ils à la fois les *indices de la direction*. Les indices de la direction sont notés $[mnp]$. La figure 1.6, *b* indique à titre d'exemple les indices des principales directions dans un cristal cubique. Les indices négatifs sont munis de signe « — » surmontant l'indice.

Si les nombres m, n, p ne sont pas des entiers, on les arrondit à des entiers les plus proches qu'on prend pour indices de la direction. Par exemple dans un réseau à faces centrées le premier nœud dans la direction de la diagonale de la face a pour coordonnées 0, 1/2, 1/2. En arrondissant ces nombres, on obtient les indices de la direction [011].

Indices des plans. Supposons qu'un plan cristallographique coupe sur les axes de coordonnées des segments m, n, p. L'équation d'un

tel plan en segments est de la forme

$$\frac{x}{m}+\frac{y}{n}+\frac{z}{p}=1, \tag{1.3}$$

où x, y, z sont les coordonnées d'un point situé dans ce plan. Si l'on réduit tous les termes de l'équation au même dénominateur D, on

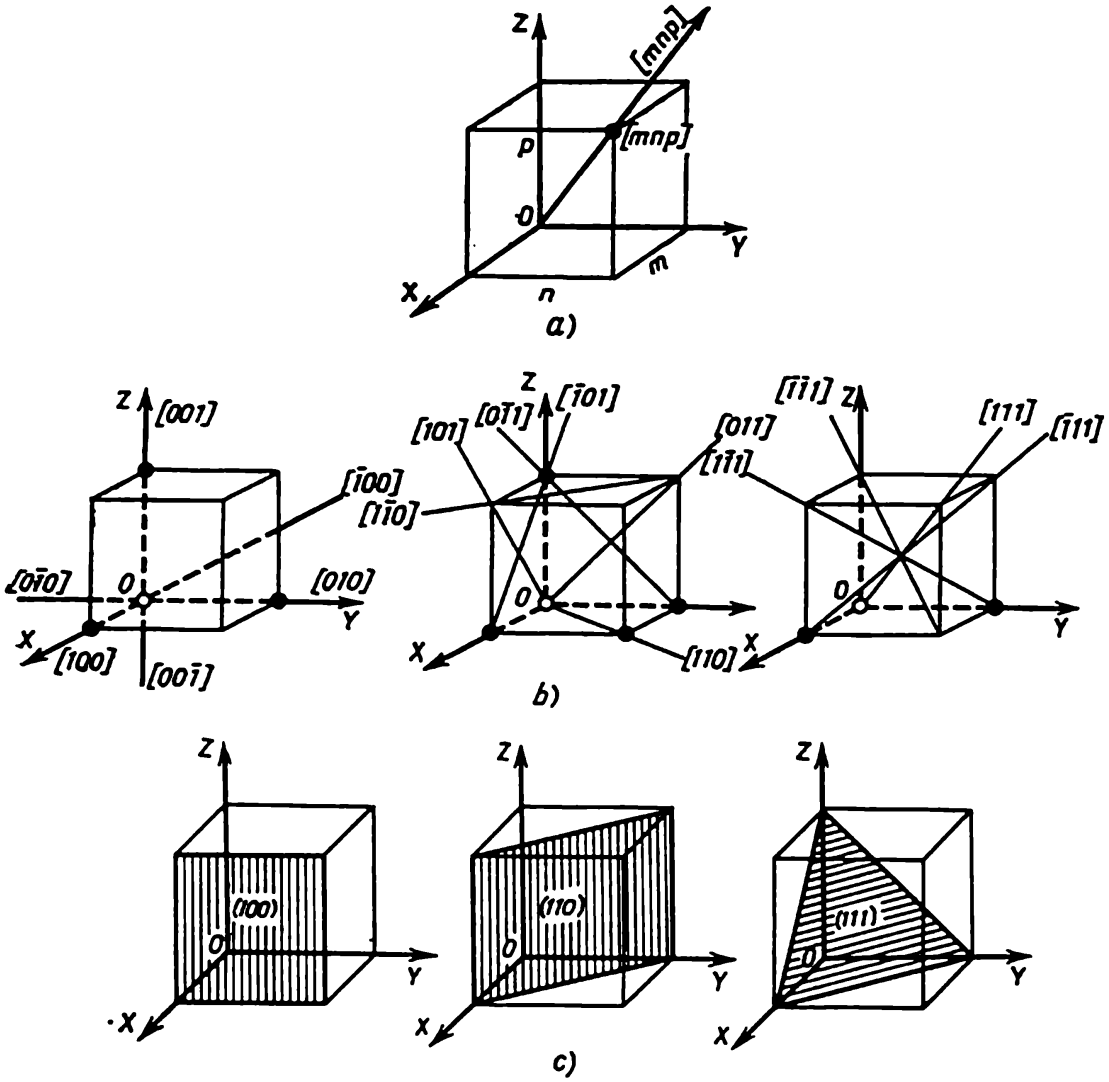

Fig. 1.6. Désignation des nœuds, des directions et des plans dans les cristaux: *a* — indices des nœuds [[*mnp*]] et des directions [*mnp*]; *b* — indices des directions principales dans un cristal cubique; *c* — indices de Miller de certains plans d'un cristal cubique.

trouve

$$hx + ky + lz = D, \tag{1.4}$$

où

$$h=\frac{D}{m},\ k=\frac{D}{n},\quad l=\frac{D}{p} \tag{1.5}$$

sont des entiers qu'on prend précisément pour *indices du plan*. Ils portent le nom d'*indices de Miller*. Le plan défini par ces indices est noté (hkl). Par exemple, si un plan coupe sur les axes de coordonnées les segments 3, 2, 1 son équation a la forme $\frac{x}{3}+\frac{y}{2}+\frac{z}{1}=1$. Après la réduction au même dénominateur on obtient $2x+3y+6z=6$. On en déduit que $h=2$, $k=3$, $l=6$. Un tel plan est noté (2 3 6).

Des plans, parallèles l'un à l'autre, pour lesquels les segments coupés sont égaux par exemple à 1, 2, 3; 2, 4, 6; 3, 6, 9, etc., sont désignés par les mêmes indices de Miller (1 2 3). Dans cet ensemble des plans parallèles on range aussi des plans à indices négatifs $(\bar{1}\,\bar{2}\,\bar{3})$. Ainsi, les indices de Miller (hkl) déterminent non pas un seul plan mais une famille de plans parallèles. Une telle famille de plans cristallographiquement et physiquement équivalents est notée $\{hkl\}$ et une famille de directions physiquement équivalentes $\langle mnp \rangle$.

Des relations (1.5) il résulte que

$$m=\frac{D}{h}, \quad n=\frac{D}{k}, \quad p=\frac{D}{l}. \tag{1.6}$$

Aussi, pour pouvoir déterminer par les indices (hkl) d'un plan les segments qu'il coupe sur les axes du réseau, est-il nécessaire de connaître les inverses des indices $(1/h,\ 1/k,\ 1/l)$ et de trouver leur dénominateur commun D. Alors $m=D/h$, $n=D/k$, $p=D/l$. Par exemple, pour un plan d'indices (1 2 3) les inverses des indices sont égaux à 1/1, 1/2, 1/3. Le dénominateur commun est égal à 6, les segments découpés sur les axes sont $m=6$, $n=3$, $p=2$.

Pour des plans parallèles aux axes de coordonnées, l'indice correspondant est nul. C'est ainsi que le plan (110) est parallèle à l'axe OZ, le plan (011) à l'axe OX, etc.

Indiquons à titre d'exemple les indices des plans principaux d'un réseau cubique (fig. 1.6, *c*). Les indices des *faces du cube* : (100), (010), (001); les indices des plans passant par les diagonales des faces (plans du dodécaèdre *rhombique*) : (110), (101), (011), etc. ; les indices des plans découpant sur les axes des segments unités (plans de l'*octaèdre*) : (111), $(1\bar{1}1)$, $(\bar{1}\bar{1}1)$, $(\bar{1}11)$, etc.

Indiquons certaines relations cristallographiques qui seront utiles pour l'exposé qui suit.

Période d'identité I. C'est la distance entre les atomes identiques le long d'une direction donnée qui se détermine comme la distance de l'origine des coordonnées au point considéré $[[mnp]]$. Dans le cas général cette période se détermine par les paramètres a, b, c, α, β, γ et par les nombres m, n, p.

Pour un réseau cubique simple

$$I=\sqrt{m^2+n^2+p^2}. \tag{1.7}$$

Comme le montre (1.7), plus petits sont les indices d'une direction, plus serrés sont les atomes le long de cette direction.

Distance (intervalle) réticulaire. Dans le cas général elle se détermine par les indices des plans (hkl) et les paramètres de la maille élémentaire. Pour les cristaux du système cubique

$$d = \frac{a}{\sqrt{h^2+k^2+l^2}}. \tag{1.8}$$

La relation (1.8) montre que plus grands sont les indices des plans, plus près l'un de l'autre sont disposés ces plans.

§ 1.3. Diffraction dans les cristaux. Réseau réciproque

Loi de Bragg-Wulff. Les méthodes les plus employées pour l'étude de la structure interne des cristaux sont des méthodes de diffraction. Elles sont fondées sur le fait que pour des ondes de longueur d'onde $\lambda \lesssim a$, où a est la constante du réseau, un cristal dont les particules constitutives se caractérisent par des arrangements hautement ordonnés constitue un réseau spatial de diffraction de qualité suffisamment bonne. Le passage d'un faisceau de telles ondes à travers le cristal a pour effet de provoquer l'apparition des figures de diffraction qui reflètent la structure interne du cristal. Comme les paramètres des réseaux des cristaux sont de l'ordre de 10^{-10} m, la longueur d'onde λ doit être du même ordre de grandeur ou moindre. Une telle longueur d'onde peuvent avoir les rayons X, les ondes de de Broglie des électrons accélérés et des neutrons lents. Ce sont eux qu'on utilise dans l'analyse radiocristallographique.

La théorie de la diffusion des rayons X par les cristaux a été construite en 1912 par Laue. Elle a joué un grand rôle aussi bien dans l'étude des propriétés des rayons X eux-mêmes que dans l'étude de la constitution des cristaux.

En 1913, Bragg et Wulff ont obtenu une relation bien simple exprimant la condition de diffraction des ondes par un système de plans atomiques parallèles (fig. 1.7). Ils ont montré que le rayon diffracté pouvait être considéré comme le résultat de la réflexion spéculaire du rayon incident par ce système de plans. Cette réflexion ne peut se produire que dans des directions θ qui satisfont à la condition suivante :

$$2d\,(hkl)\cdot \sin\theta = n\lambda, \tag{1.9}$$

où $d\,(hkl)$ est la distance entre les plans atomiques voisins (hkl), λ la longueur d'onde, n un entier indiquant l'ordre de diffraction. La relation (1.9) est appelée formule ou *loi* de Bragg-Wulff.

Réseau réciproque. Une forme commode pour l'utilisation pratique des conditions de diffraction par les cristaux a été propo-

sée par Ewald qui a introduit à cet effet la notion de réseau réciproque.

Le réseau réciproque est l'image du réseau ordinaire dans l'espace inverse ou dans l'espace des vecteurs d'onde. Chaque nœud du réseau

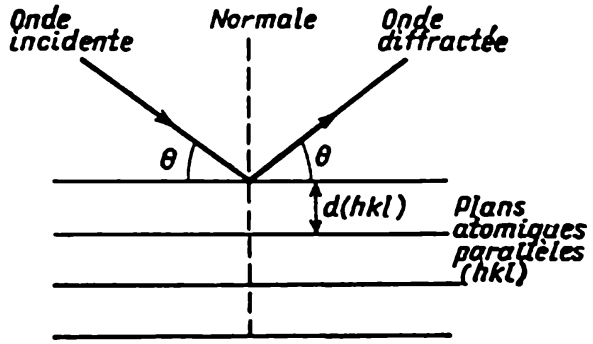

Fig. 1.7. Loi de Bragg-Wulff.
L'onde diffractée peut être considérée comme une réflexion spéculaire de l'onde incidente par un système de plans atomiques parallèles dans la direction de θ satisfaisant à la loi de Bragg-Wulff.

réciproque représente un plan de diffraction dans le cristal et a les mêmes indices que ce plan. Par plan de diffraction on entend non seulement des plans réels d'indices (*hkl*) passant par les nœuds du réseau cristallin mais également des plans ayant des indices multiples des indices de Miller : (*nh nk nl*), où *n* est un entier. A la valeur

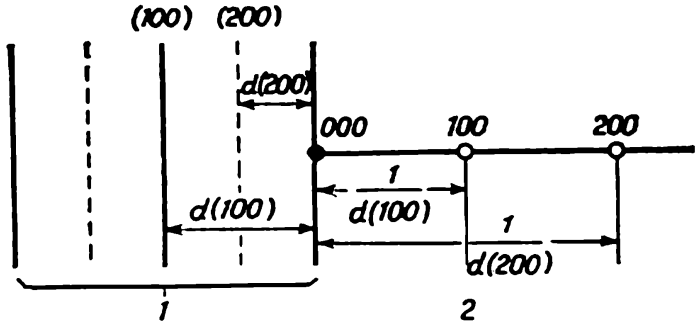

Fig. 1.8. Plans du réseau cristallin et nœuds correspondants du réseau réciproque : *1* — plans du cristal ; *2* — nœuds du réseau réciproque (Ch. S. Barret, *Structure des métaux*, Moscou, Metallourgizdat, 1948).

$n = 1$ correspond la diffraction du premier ordre, à la valeur $n = 2$ la diffraction du deuxième ordre et ainsi de suite. Les nœuds du réseau réciproque se situent sur une ligne perpendiculaire au système de plans donnés dans le cristal, à des distances de l'origine des coordonnées égales aux inverses des distances réticulaires de ces plans.

La figure 1.8 montre à titre d'exemple les plans (100) et (200) du cristal et les nœuds du réseau réciproque qui leur correspondent 100, 200 ; 000 est l'origine des coordonnées du réseau réciproque.

En représentant de cette façon toutes les autres familles de plans du cristal, on construit un réseau réciproque tridimensionnel dont les nœuds représentent tous les plans de diffraction réels et imaginaires. La figure 1.9 montre la construction des nœuds du réseau

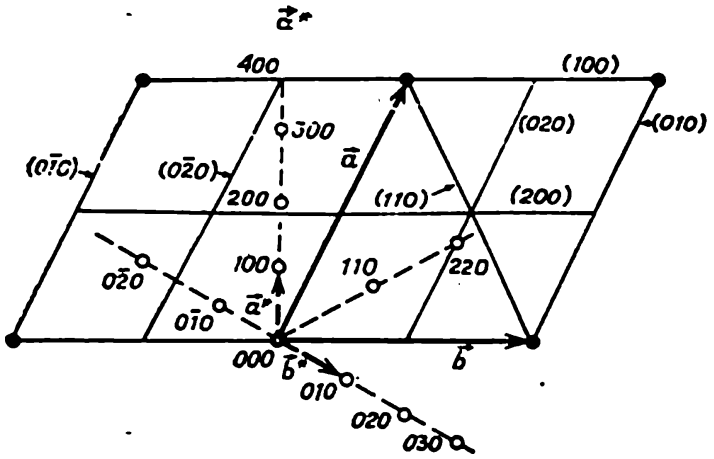

Fig. 1.9. Partie d'une couche:
○ — du réseau réciproque; ● — du réseau cristallin d'un cristal monoclinique (Ch. S. Barret, *Structure des métaux*, Moscou, Metallourgizdat, 1948).

réciproque pour une partie de la couche d'un cristal monoclinique et la figure 1.10 représente le réseau réciproque d'un cristal orthorhombique.

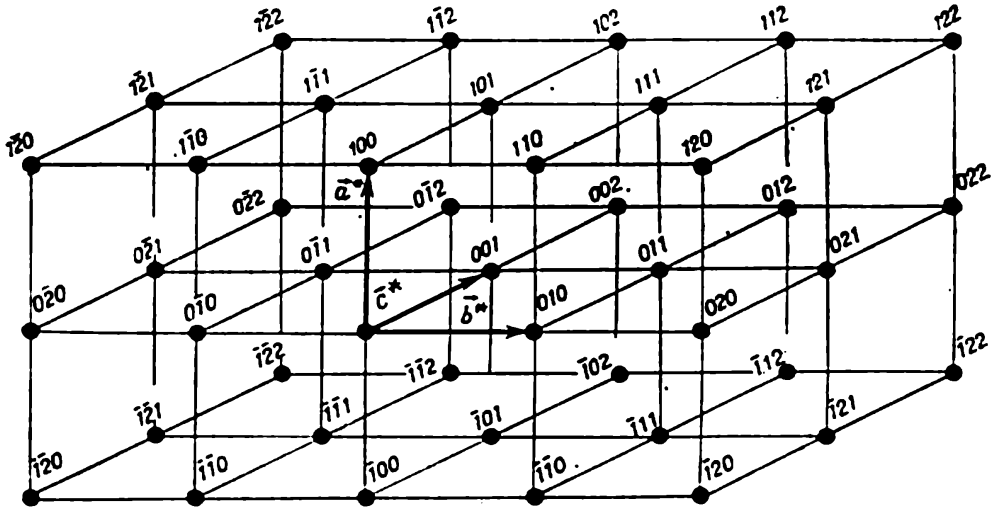

Fig. 1.10 Réseau réciproque d'un cristal orthorhombique (Ch. S. Barret, *Structure des métaux*, Moscou, Metallourgizdat, 1948).

Les nœuds du réseau réciproque forment toujours un réseau spatial simple, quel que soit le type du réseau spatial du cristal. Ils se situent aux sommets des mailles élémentaires dont les vecteurs **a***, **b***, **c*** sont liés aux vecteurs **a**, **b**, **c** de la maille élémentaire du

cristal par les relations suivantes:

$$\mathbf{a}^* = \frac{1}{V}[\mathbf{bc}], \quad \mathbf{b}^* = \frac{1}{V}[\mathbf{ca}], \quad \mathbf{c}^* = \frac{1}{V}[\mathbf{ab}], \tag{1.10}$$

où V est le volume de la maille élémentaire du réseau direct. Ici, $[\mathbf{bc}]$, $[\mathbf{ca}]$, $[\mathbf{ab}]$ sont les produits vectoriels des vecteurs élémentaires du réseau direct. Par conséquent, $\mathbf{a}^* \perp [\mathbf{bc}]$, $\mathbf{b}^* \perp [\mathbf{ca}]$, $\mathbf{c}^* \perp [\mathbf{ab}]$. Les modules des vecteurs élémentaires du réseau réciproque ont pour valeur

$$|\mathbf{a}^*| = \frac{1}{d(100)}, \quad |\mathbf{b}^*| = \frac{1}{d(010)}, \quad |\mathbf{c}^*| = \frac{1}{d(001)}, \tag{1.11}$$

et son volume est égal à

$$V^* = [\mathbf{a}^*\mathbf{b}^*]\,\mathbf{c}^* = \frac{1}{V},$$

de sorte que

$$V^* V = 1.$$

Le vecteur de translation dans le réseau réciproque a pour expression

$$\mathbf{r}^*(hkl) = h\mathbf{a}^* + k\mathbf{b}^* + l\mathbf{c}^*. \tag{1.12}$$

Il est perpendiculaire au plan (hkl) du cristal et son module est égal à l'inverse de la distance réticulaire

$$|\mathbf{r}^*(hkl)| = \frac{1}{d(hkl)}. \tag{1.13}$$

On l'appelle ordinairement *vecteur de réseau réciproque*.

Considérons maintenant les *conditions de diffraction*, en utilisant, avec Ewald, la notion de réseau réciproque.

Menons vers l'origine des coordonnées O du réseau réciproque le vecteur $\mathbf{S}_{inc}$ du rayon incident de longueur $|\mathbf{S}_{inc}| = |AO| = 1/\lambda$ et décrivons par ce vecteur une sphère de centre au point A (fig. 1.11). Alors les directions de diffraction possibles correspondent aux points d'intersection de la sphère construite avec les nœuds du réseau réciproque. Notamment le maximum de diffraction pour le rayon $\mathbf{S}_{inc}$ est observé dans la direction AB qui fait avec la direction de $\mathbf{S}_{inc}$ un angle 2θ. En effet, après la réflexion sur les plans (hkl) le rayon se propagera dans la direction de $\mathbf{S}_{ref}$ qui fait avec le plan (hkl) l'angle θ. En déplaçant ce vecteur parallèlement à lui-même de façon que son origine coïncide avec le point A et en reliant entre elles les extrémités des vecteurs $\mathbf{S}_{inc}$ et $\mathbf{S}_{ref}$, on obtient un vecteur $\mathbf{S}$ qui caractérise la diffusion du rayon dans la réflexion. Le vecteur $\mathbf{S}$ est appelé *vecteur de diffusion*. Comme il est visible sur la figure 1.11, son module est égal au module du vecteur du réseau réciproque: $|\mathbf{S}| = \frac{1}{d(hkl)}$ correspondant au système de plans $\{hkl\}$. D'un autre

côté, l'examen de la figure 1.11 montre que $|S| = 2|S_{inc}| \sin\theta = 2\cdot\frac{1}{\lambda}\sin\theta = \frac{1}{d(hkl)}$. On en déduit

$$2d(hkl)\sin\theta = \lambda.$$

Cette expression est celle de la loi de Bragg-Wulff. Si h, k, l contiennent un multiplicateur commun n, alors, suivant (1.8), la distance entre les plans $(nh\ nk\ nl)$ est n fois plus faible qu'entre les plans

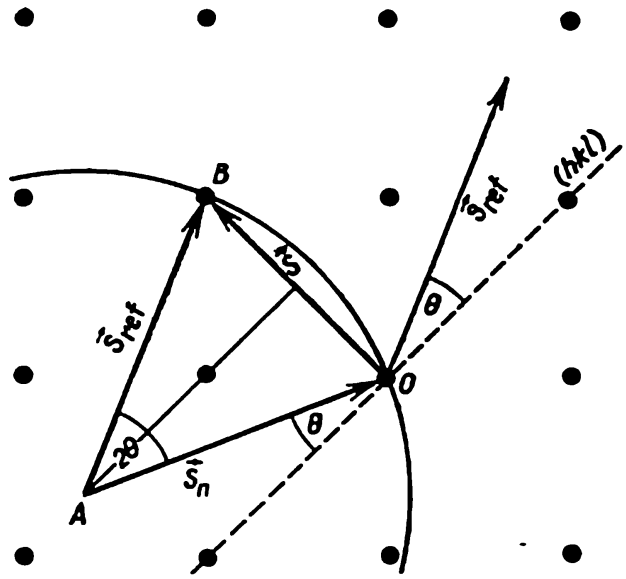

Fig. 1.11. Construction d'Ewald dans le réseau réciproque.

(hkl) et le vecteur du réseau réciproque est n fois plus grand que le vecteur $\frac{1}{d(hkl)}$. Par conséquent, $2\frac{1}{\lambda}\sin\theta = \frac{n}{d(hkl)}$. On en tire

$$2d(hkl)\sin\theta = n\lambda,$$

ce qui coïncide avec (1.9). Ainsi le faisceau de diffraction peut être considéré soit comme une réflexion du n-ième ordre sur des plans du réseau d'indices de Miller minimaux (hkl), soit comme une réflexion du premier ordre sur un système de plans du cristal séparés l'un de l'autre par des distances qui sont n fois plus petites que les distances entre les plans (hkl).

Il est commode de porter dans la direction du rayon incident son vecteur d'onde $\mathbf{k}$ de module $|\mathbf{k}| = 2\pi/\lambda$ et dans la direction du rayon réfléchi son vecteur d'onde $\mathbf{k}'$ dont le module, lors de la diffusion élastique, est lui aussi égal à $2\pi/\lambda$. Alors

$$\mathbf{k}' - \mathbf{k} = \mathbf{G} \text{ ou } \mathbf{k}' = \mathbf{k} + \mathbf{G}, \tag{1.14}$$

où $\mathbf{G} = 2\pi\mathbf{S}$. Le vecteur $\mathbf{G}$ est appelé lui aussi *vecteur du réseau réciproque*. Son module est

$$|\mathbf{G}| = 2\pi/d\,(hkl). \tag{1.15}$$

En élevant au carré les deux membres de la relation (1.14), on obtient $k'^2 = (\mathbf{k} + \mathbf{G})^2$. Comme $k'^2 = k^2$, on a

$$k^2 = (\mathbf{k} + \mathbf{G})^2 \tag{1.16}$$

ou

$$2\mathbf{kG} + G^2 = 0. \tag{1.17}$$

Les relations (1.16) et (1.17) sont équivalentes à la condition de Bragg de réflexion des ondes dans le cristal. C'est pourquoi on les considère comme une autre expression de la loi de Bragg-Wulff.

Etant donné que l'impulsion du photon

$$\mathbf{p}_{ph} = \hbar\mathbf{k}, \tag{1.18}$$

l'espace inverse du réseau direct sera également l'espace des impulsions. En multipliant (1.14) par $\hbar$, on obtient la formule suivante:

$$\mathbf{p}' = \mathbf{p} + \hbar\mathbf{G}. \tag{1.19}$$

Elle exprime la variation de l'impulsion du photon par suite de la diffusion élastique.

§ 1.4. Méthodes expérimentales de détermination de la structure des cristaux

De la loi de Bragg-Wulff il résulte que les figures de diffraction qui apparaissent lorsque les rayons X passent par un cristal, sont directement liées à la structure de ce cristal. C'est sur ce fait que sont fondées les méthodes de diffraction d'étude de la structure des cristaux. Dans ces méthodes, au lieu des rayons X on peut employer des faisceaux d'électrons accélérés, de neutrons lents et d'autres particules.

Examinons sommairement les principales méthodes d'analyse structurelle des cristaux.

Méthode de Laue. Dans cette méthode un faisceau étroit de rayons X «blancs» (polychromatiques) contenant toutes les longueurs d'ondes dans un intervalle $\Delta\lambda$ suffisamment large est dirigé sur un échantillon monocristallin K fixé dans un support spécial (fig. 1.12). De tous les rayons constituant ce faisceau ne subiront la diffraction que ceux de longueur d'onde λ qui tombent sur des plans atomiques correspondants MM sous un angle θ satisfaisant à la loi de Bragg-Wulff. Sur la pellicule radiographique A on voit apparaître des maximums de diffraction pour des rayons ayant traversé le cristal (prise de vue directe) ou réfléchis sur le cristal dans le sens inverse et tombant sur la pellicule B (prise de vue inverse). Le diagramme de Laue,

dont un exemple est montré sur la figure 1.13, est constitué par une série de taches disposées suivant une loi déterminée et caractérisant la symétrie du cristal. Si le faisceau incident passe dans la

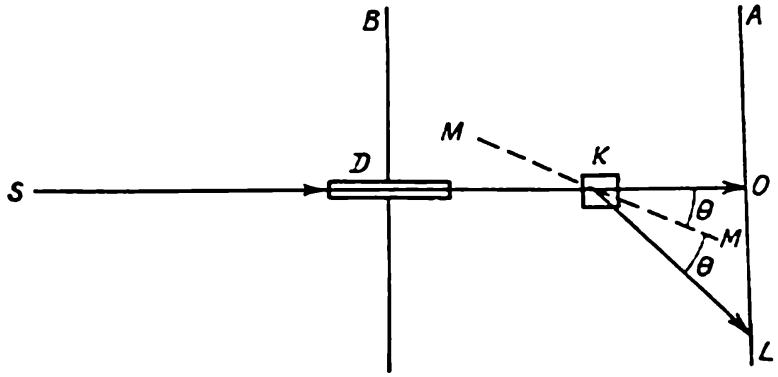

Fig. 1.12. Méthode de Laue:
SO — faisceau primaire de rayons X « blancs »; *D* — collimateur; *K* — cristal ; *MM* — trace du plan cristallographique ; θ — angle de réflexion de Bragg; *KL* — rayon « réfléchi »; *A*, *B* — pellicules radiographiques ; *O* — centre du diagramme radiographique.

direction de l'axe du 4-ième ordre par exemple, le diagramme de Laue possède un axe de symétrie du 4-ième ordre. Le diagramme de Laue représenté sur la figure 1.13 a été obtenu en faisant passer un

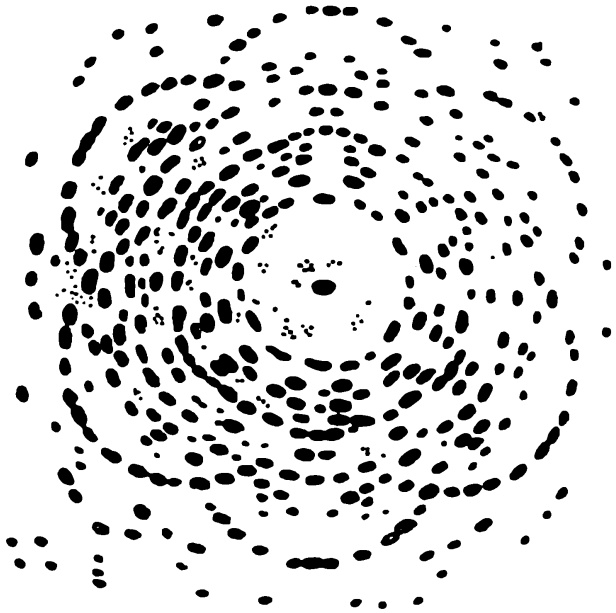

Fig. 1.13. Diagramme de Laue d'un cristal de quartz.

faisceau de rayons à travers un cristal de quartz le long de l'axe de symétrie du 6-ième ordre.

La méthode de Laue est largement utilisée pour la détermination de l'orientation des cristaux et rarement pour l'étude de leur struc-

ture parce que le grand nombre de réflexions sur un même plan atomique a pour résultat une superposition des taches obtenues par suite de la réflexion de divers ordres, ce qui rend difficile l'interprétation des taches et la détermination de la structure du cristal.

Méthode du cristal tournant. Le monocristal *K* est fixé (fig. 1.14) dans un support spécial suivant l'axe d'une chambre cylindrique dont les parois sont recouvertes de pellicule radiographique. L'échantillon est illuminé avec un faisceau étroit de rayons X monochromatiques. La rotation de l'échantillon autour de l'axe fait varier

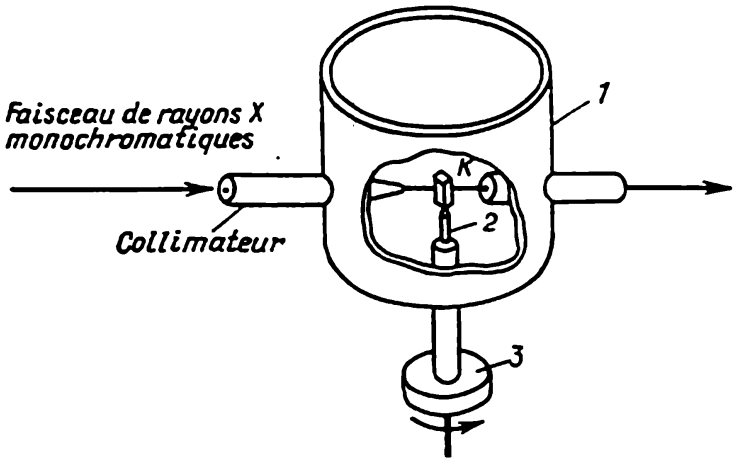

Fig. 1.14. Méthode du cristal tournant:
K — monocristal; *1* — chambre cylindrique dont la surface intérieure porte une pellicule radiographique; *2* — support de cristal; *3* — poulie pour l'entraînement en rotation du cristal (du livre de Ch. S. Barret, *Structure des métaux*).

l'angle θ entre les rayons et un plan atomique donné. Si cet angle prend une valeur telle que pour la longueur d'onde donnée λ est satisfaite la condition de Bragg-Wulff, il apparaît un maximum de diffraction qui est fixé sur la pellicule. Il est évident que des plans atomiques verticaux donneront des maximums de diffraction situés dans un plan horizontal.

Une variété de méthode du cristal tournant est la *méthode du monocristal balançant* dans laquelle le cristal tourne autour de son axe dans un intervalle déterminé d'angles. Cela diminue la probabilité de superposition des maximums de diffraction dus à divers plans. Si on fait tourner aussi la pellicule en synchronisme avec la rotation du cristal, on peut exclure presque totalement les superpositions des taches. Pour l'enregistrement des maximums de diffraction on utilise, dans certains spectromètres, divers compteurs permettant de déterminer non seulement les positions mais également l'intensité des maximums de diffraction.

Méthode des poudres (méthode de Debye). Pour l'enregistrement de radiogrammes par la méthode des poudres un pinceau monochromatique de rayons X est dirigé sur une poudre cristalline (fig. 1.15).

Si les grains de cette poudre sont suffisamment fins et orientés de façon désordonnée, il existe toujours parmi eux des grains dont les plans réticulaires forment avec le rayon incident des angles de Bragg. Dans des grains différents de tels angles seront formés par des plans

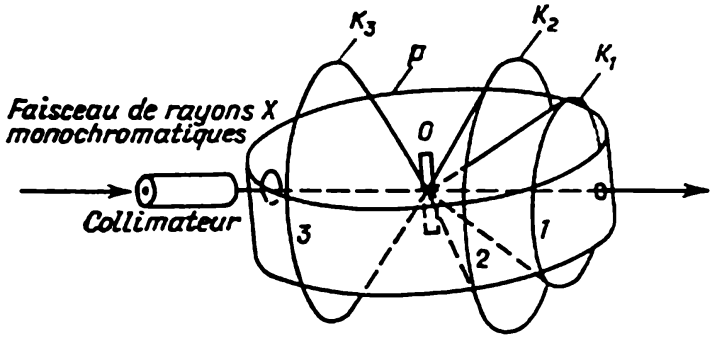

Fig. 1.15. Méthode des poudres:

O — échantillon de poudre ou de polycristal à grains fins; *P* — pellicule placée dans une chambre cylindrique; K_1, K_2 — cônes coaxiaux suivant lesquels les rayons réfléchis sous des angles de Bragg se propagent dans le sens direct; K_3 — cônes suivant lesquels les rayons réfléchis sous des angles de Bragg se propagent dans le sens inverse; *1*, *2*, *3* — lignes d'intersection des cônes avec la pellicule cylindrique.

réticulaires ayant, dans le cas général, des indices (*hkl*) différents. On observe donc simultanément toutes les réflexions qu'on pourrait obtenir pour un monocristal de la même substance lors de sa rotation. Du fait de la symétrie totale des réflexions possibles par rapport

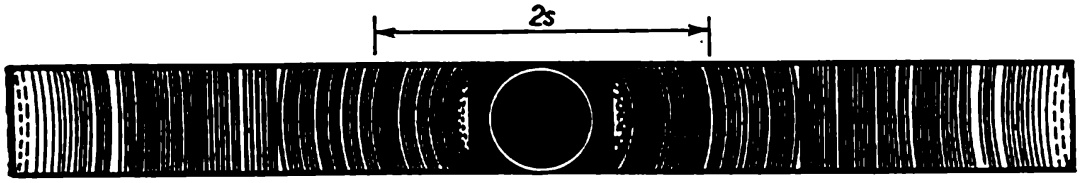

Fig. 1.16. Diagramme de Debye d'une poudre cristalline.

au pinceau primaire les rayons « réfléchis » sont dirigés le long d'une famille de cônes coaxiaux dont l'axe est la direction du pinceau primaire et les angles d'ouverture sont égaux à 4θ, où θ est l'angle de réflexion de Bragg.

L'échantillon est fixé au centre d'une chambre cylindrique dont la surface intérieure porte une pellicule photographique. Les cônes se coupent avec la pellicule suivant des courbes du 4-ième ordre dont un exemple est schématiquement représenté sur la figure 1.16. En mesurant la distance ($2s$) entre les lignes, on peut déterminer les angles de réflexion sur des plans réticulaires différents et donc calculer, à l'aide de l'équation de Bragg-Wulff, les distances réticulaires correspondantes.

§ 1.5. Classification des cristaux suivant les types de liaison

La structure cristalline et les propriétés des corps solides se déterminent pour une large part par le caractère des forces de liaison entre les particules. Etudions sommairement les principaux types de liaison dans les solides et les structures cristallines types qui leur correspondent.

Cristaux ioniques. Les atomes qui occupent dans le système périodique de Mendéléev des places près des gaz inertes peuvent assez facilement prendre la configuration de leurs couches électroniques en cédant ou en acceptant des électrons. C'est ainsi que dans les atomes de métaux alcalins qui suivent immédiatement les gaz inertes, l'électron de valence gravite en dehors de la couche remplie et est lié au noyau de façon relativement faible si bien que l'énergie d'ionisation de ces atomes est faible (de l'ordre de 5 eV). Chez les halogènes qui précèdent immédiatement les gaz inertes, il manque un électron pour former une couche électronique stable de gaz noble. Ils possèdent donc une forte affinité électronique pour l'électron additionnel (de l'ordre de 3 eV).

La liaison entre les atomes de ce genre, c'est-à-dire entre les métaux forts et les halogènes, s'effectue de la façon suivante. Tout d'abord il se produit le changement de charge des deux atomes: l'électron de l'atome de métal passe dans l'atome d'halogène. Il en résulte que l'atome de métal se transforme en ion chargé positivement (cation) et l'atome d'halogène en ion chargé négativement (anion). Ces ions interagissent l'un avec l'autre suivant la loi de Coulomb comme deux charges de noms contraires. Une telle liaison est dite *ionique* ou *hétéropolaire.*

L'énergie d'attraction des ions distants l'un de l'autre de r est égale à $U_{att} = -e^2/(4\pi\varepsilon_0 r)$, où e est la charge de l'ion. Lorsque les ions se rapprochent à une distance suffisamment petite, il se crée la répulsion dont l'énergie est suivant Born égale à $U_{rép} = B/r^n$, où B et n sont des constantes. L'énergie d'interaction résultante a pour valeur

$$U = \frac{B}{r^n} - \frac{e^2}{4\pi\varepsilon_0 r}. \tag{1.20}$$

La distance d'équilibre r_0 entre les ions s'obtient si on annule la dérivée de U par rapport à r (fig. 1.1):

$$r_0 = \left(\frac{4\pi\varepsilon_0 nB}{e^2}\right)^{\frac{1}{n-1}}.$$

En déterminant à partir de cette dernière expression B et en portant sa valeur dans (1.20), on obtient

$$U = -U_l = -\frac{e^2}{4\pi\varepsilon_0 r_0}\left(1-\frac{1}{n}\right). \qquad (1.21)$$

L'énergie d'un réseau construit à partir de N paires d'ions a pour expression

$$U_{rés} = -NA\frac{e^2}{4\pi\varepsilon_0 r_0}\left(1-\frac{1}{n}\right), \qquad (1.22)$$

où A est la *constante de Madelung* qui tient compte du fait que l'ion considéré interagit non seulement avec son voisin le plus proche mais également avec les autres ions du réseau.

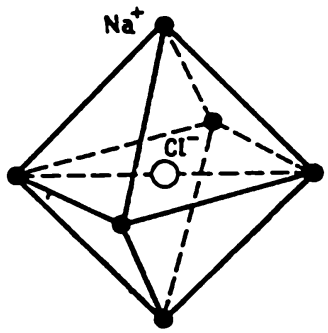

Fig. 1.17. Réseau cubique du chlorure de sodium de nombre de coordination $z = 6$.

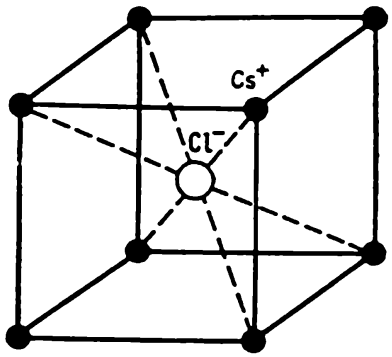

Fig. 1.18. Réseau cubique de chlorure de césium de nombre de coordination $z = 8$.

Dans les cristaux ioniques il n'existe pas de molécules distinctes: chaque ion chargé positivement est entouré d'un nombre correspondant d'ions chargés négativement et il en est exactement de même pour les ions chargés négativement dont chacun est entouré par un même nombre d'ions chargés positivement. Le nombre z de voisins les plus proches entourant un ion (atome) est appelé *nombre de coordination*. Pour les réseaux ioniques ce nombre dépend du rapport du rayon r_c du cation au rayon r_a de l'anion, c'est-à-dire de $R = r_c/r_a$. Pour $R > 0{,}73$, ce qui a lieu par exemple pour CsCl, $z = 8$; pour $0{,}41 < R < 0{,}73$, comme chez NaCl, $z = 6$; pour $R < 0{,}41$, comme chez LiI, $z = 4$.

La figure 1.17 montre un réseau cubique du cristal de NaCl dans lequel chaque ion chlore est entouré de six ions Na^+ et chaque ion sodium de six ions chlore Cl^-. La figure 1.18 montre la disposition de huit ions césium autour d'un ion chlore dans le réseau de CsCl.

La liaison ionique est une liaison forte dont l'énergie est comprise entre quelques centaines et quelques milliers de kilojoules par

mole. Cela explique pourquoi les solides à liaison ionique possèdent une résistance et une dureté considérables, un point de fusion élevé et une forte chaleur de sublimation. Leur conductibilité électrique a une nature ionique et dépend fortement de la température. De nombreux cristaux ioniques sont transparents dans une large gamme de longueurs d'onde.

Cristaux atomiques (covalents). Pour illustrer la nature de la liaison covalente, considérons un exemple bien simple d'interaction

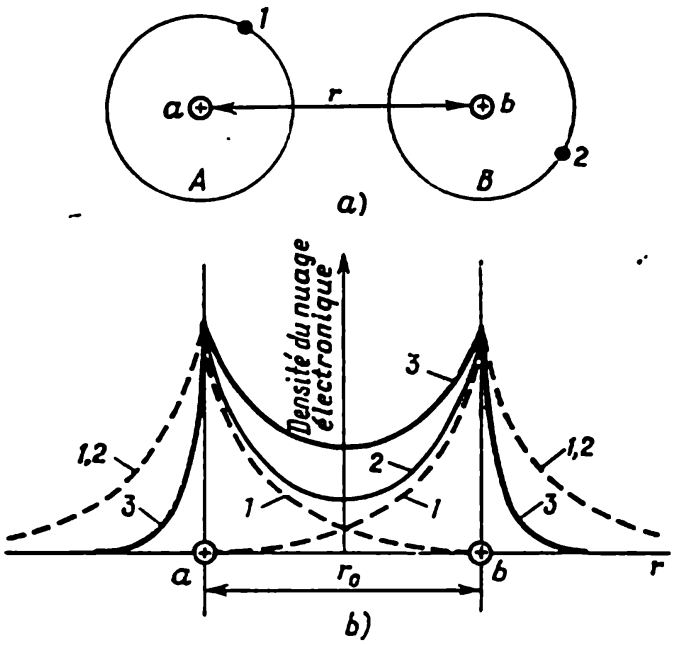

Fig. 1.19. Explication de la nature de la liaison covalente dans une molécule d'hydrogène :

a — représentation schématique des atomes d'hydrogène ; *b* — répartition de la densité électronique dans le système de deux atomes d'hydrogène ; *1* — répartition de la densité électronique dans des atomes d'hydrogène isolés ; *2* — densité électronique qui serait obtenue par une simple superposition des nuages électroniques des atomes isolés rapprochés à une distance r_0 ; *3* — répartition réelle de la densité électronique dans une molécule d'hydrogène.

de deux atomes d'hydrogène *A* et *B* (fig. 1.19, *a*). Comme la densité du nuage électronique décrivant l'état d'un électron dans l'atome décroît rapidement à mesure qu'on s'éloigne du noyau, à une grande distance entre les atomes la probabilité de localiser l'électron *1* près du noyau *b* et l'électron *2* près du noyau *a* est négligeable. Aussi, les atomes *A* et *B* peuvent-ils être considérés comme isolés et l'énergie du système constitué de deux atomes comme égale à $2E_0$, où E_0 est l'énergie d'un atome isolé.

Au fur et à mesure que les atomes se rapprochent, la probabilité de passage des électrons dans les noyaux « étrangers » augmente. Pour $r \approx 0{,}2$ nm il se produit un recouvrement notable des nuages électroniques de ces atomes et la fréquence de passage atteint une valeur de $\sim 10^{14}$ s^{-1}. Si les atomes se rapprochent encore, le degré de

recouvrement des nuages électroniques augmente et la fréquence d'échange d'électrons entre les atomes devient si élevée que l'affirmation que l'électron *1* appartient à l'atome *A* et l'électron *2* à l'atome *B* cesse d'avoir un sens. Cela correspond à la création d'un nouvel état qui n'est pas inhérent à un système constitué de deux atomes isolés. Une chose remarquable est que dans cet état les électrons appartiennent aux deux noyaux à la fois ou, comme on dit, sont *mis en commun*.

La mise en commun des électrons s'accompagne d'une redistribution de la densité électronique et d'une variation de l'énergie du système par rapport à l'énergie totale $2E_0$ des atomes isolés. La redistribution consiste à « entraîner » les nuages électroniques dans l'espace entre les noyaux : à une petite distance du noyau, en dehors de cet espace, la densité des nuages diminue par rapport à la densité des nuages dans les atomes isolés tandis que dans l'espace entre les noyaux cette densité augmente par rapport à celle qui pourrait être obtenue par l'addition dans cet espace des nuages des atomes isolés (fig. 1.19, *b*).

L'apparition d'un état à densité élevée du nuage électronique, remplissant l'espace entre les noyaux, provoque toujours une diminution de l'énergie du système et conduit à l'apparition de forces d'attraction. De façon imagée, tout se passe comme si le nuage électronique de densité élevée, qui est formé dans l'espace entre les noyaux par deux électrons mis en commun, attirait les noyaux l'un vers l'autre en cherchant à les rapprocher au maximum l'un de l'autre.

Un calcul de mécanique quantique montre que le passage des atomes isolés H à une molécule H_2 donne lieu aux phénomènes suivants :

— l'énergie d'attraction des électrons par les noyaux varie de —4260 kJ/mole ;

— il apparaît l'énergie de répulsion des électrons égale à +1540 kJ/mole ;

— il apparaît l'énergie de répulsion des noyaux égale à +1905 kJ/mole ;

— l'énergie cinétique des électrons augmente de +402 kJ/mole.

L'énergie résultante doit ainsi être égale à $U_l = -413$ kJ/mole, ce qui est proche de la valeur expérimentale égale à —456 kJ/mole.

L'énergie de liaison qui prend naissance par suite de la mise en commun des électrons deux à deux est souvent appelée énergie *d'échange* parce qu'elle apparaît par suite d'un échange d'électrons entre les atomes. En réalité, c'est une énergie électrostatique d'interaction entre le nuage électronique de densité accrue formé entre les atomes et les noyaux de ces atomes. Elle peut être donnée par l'expression approchée suivante :

$$U_{éch} = -2J\,(\mathbf{S}_1 \cdot \mathbf{S}_2), \qquad (1.23)$$

où S_1, S_2 sont les spins des électrons mis en commun; J est une intégrale dite *d'échange* qui exprime le résultat de la superposition des nuages électroniques des atomes rapprochés. Le signe de J dans les molécules est en règle générale négatif. Pour que $U_{éch}$ soit dans ce cas négative, c'est-à-dire pour qu'entre les atomes s'exerce une force d'attraction, les spins des électrons à apparier (à mettre en commun) doivent être *antiparallèles*. Lorsque les spins sont parallèles, $U_{éch} > 0$ et les atomes se repoussent du fait que suivant le principe de Pauli la densité du nuage électronique dans l'espace entre les noyaux se trouve voisine de zéro.

Les propriétés caractéristiques de la liaison covalente sont sa *saturabilité* et sa *directivité*. La saturabilité consiste en ce que chaque atome est capable de former une liaison covalente seulement avec un nombre bien déterminé de voisins (cela détermine le nombre de coordination du réseau dans lequel s'exerce cette liaison). La liaison covalente se forme dans la direction où est disposée la plus grande partie du nuage électronique correspondant aux électrons de valence. Il en résulte un recouvrement maximal des nuages des électrons à coupler. C'est là que réside le caractère directif de la liaison covalente qui joue un grand rôle dans la détermination de la disposition spatiale des atomes dans le réseau.

Un exemple typique de structures cristallines à liaison covalente est fourni par la structure du type de diamant dont le réseau est représenté par la figure 1.5, *b*. D'autres éléments chimiques du sous-groupe IVB du tableau de Mendéléev — le silicium, le germanium, l'étain gris — ont le même réseau. Les atomes de ces éléments ont quatre électrons de valence qu'ils mettent en commun avec les électrons de quatre atomes voisins comme il est schématiquement représenté sur la figure 1.20. Aussi, chaque atome dans le réseau est-il entouré de quatre voisins les plus proches situés aux sommets d'un tétraèdre représenté en pointillé sur la figure 1.5, *b*.

L'énergie d'une liaison covalente est bien élevée, de l'ordre de quelques centaines de kilojoules par mole. Aussi les cristaux covalents se caractérisent-ils par une haute résistance mécanique et la dureté, possèdent les points de fusion élevés et une forte chaleur de sublimation. Au point de vue électrique ils sont des diélectriques au voisinage de 0 K et des semiconducteurs pour $T > 0$ K. Plusieurs d'entre eux sont transparents dans la partie ondes longues du spectre.

Cristaux métalliques. Les électrons de valence périphériques dans les atomes de métaux sont liés de façon relativement faible. Lors de la formation de l'état solide, les atomes se disposent si près l'un de l'autre que les électrons de valence acquièrent la faculté de quitter leurs atomes et de se déplacer librement à l'intérieur du réseau cristallin. Une telle mise en commun des électrons conduit à leur répartition uniforme dans le réseau du métal. La liaison prend naissance par suite de l'interaction des ions positifs du réseau avec le

gaz électronique : les électrons se trouvant entre les ions attirent ces derniers les uns vers les autres en équilibrant les forces de répulsion qui s'exercent entre les ions eux-mêmes.

La liaison métallique est propre à des métaux typiques et à de nombreux composés intermétalliques. L'énergie de cette liaison est elle aussi bien élevée (quelques centaines de kilojoules par mole). Les réseaux caractéristiques des métaux sont des réseaux à faces centrées cubiques, cubiques centrés et un réseau hexagonal compact.

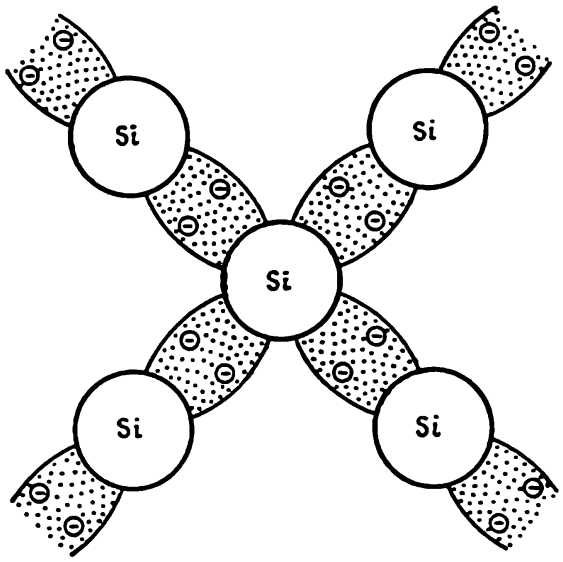

Fig. 1.20. Formation d'une liaison covalente dans le silicium par la mise en commun, deux à deux, des électrons.

Les métaux possèdent des hautes conductibilités électrique et thermique dues à la présence d'électrons libres et une haute plasticité due au caractère de la liaison permettant le déplacement sans destruction d'une partie du réseau par rapport à l'autre à des distances considérables. La présence d'électrons libres explique aussi la non-transparence et un haut pouvoir réflecteur des métaux.

Cristaux moléculaires. La forme la plus générale de la liaison qui apparaît entre les atomes et les molécules de toute espèce est la liaison de Van der Waals. Elle se manifeste presque dans toute sa pureté entre des molécules à liaisons chimiques saturées (O_2, H_2, CH_4 et autres), ainsi qu'entre les atomes de gaz inertes. Dans le cas général cette liaison contient des interactions de dispersion, d'orientation et d'induction.

Interaction de dispersion. Considérons l'exemple le plus simple de l'interaction de deux atomes d'hélium (fig. 1.21). La distribution de la densité électronique dans l'atome d'hélium présente une symétrie sphérique de sorte que son moment électrique est nul. Lorsque les atomes se rapprochent, ils se polarisent et dans le mouvement des

électrons il s'établit une corrélation qui provoque l'apparition de forces d'interaction. Si le mouvement est corrélé comme le montre la figure 1.21, *a*, il se produit entre les dipôles instantanés une attraction qui fait naître entre les atomes des forces de liaison ; pour la corrélation montrée sur la figure 1.21, *b* c'est la répulsion qui se produit entre les atomes. Comme la réalisation de la corrélation représentée sur la figure 1.21, *a* entraîne une diminution de l'énergie du

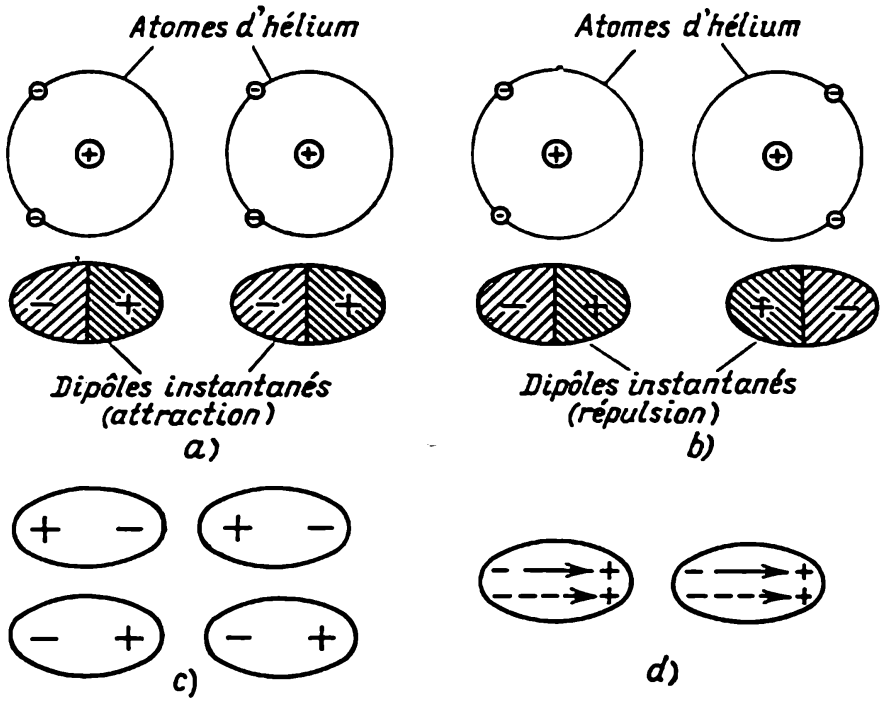

Fig. 1.21. Liaison de Van der Waals:

a, *b* — interaction de dispersion de London de deux atomes d'hélium: attraction instantanée (*a*) et répulsion (*b*) des dipôles électriques instantanés qui prennent naissance par suite de la polarisation mutuelle des atomes (du livre de G. Pimentel, R. Spratly, *Comment la mécanique quantique explique-t-elle la liaison chimique* ?) ; *c* — interaction d'orientation des molécules possédant des moments électriques constants (pour l'explication, voir le texte) ; *d* — interaction d'induction des molécules polaires possédant une autre polarisabilité (pour l'explication, voir le texte).

système, cette configuration est la plus probable et est le plus souvent réalisée, ce qui détermine précisément l'apparition d'une force d'attraction s'exerçant en permanence. Une telle interaction qui a été étudiée quantiquement pour la première fois par London est appelée interaction *de dispersion*. L'énergie de cette interaction a pour valeur

$$U_{disp} = -\frac{3}{4}\frac{\alpha^2 I}{r^6},$$

où r est la distance entre les molécules ; α leur polarisabilité ; I l'énergie d'excitation des molécules.

Interaction d'orientation. Si les molécules possèdent un moment dipolaire constant M, c'est-à-dire sont polaires, il se produit entre

elles une interaction électrostatique qui tend à ordonner la disposition des dipôles (fig. 1.21, *c*). Une telle mise en ordre provoque une diminution de l'énergie du système et l'apparition de forces d'attraction. Aux températures élevées l'énergie *d'interaction d'orientation* a pour valeur

$$U_{or} = \frac{M^4}{24\pi^2\varepsilon_0^2 kT} \frac{1}{r^6},$$

où r est la distance entre les molécules.

Interaction d'induction. Dans les molécules polaires possédant une haute polarisabilité peut apparaître, sous l'effet des dipôles constants des molécules voisines, un moment induit (la figure 1.21, *d* montre les dipôles induits en pointillé). L'énergie d'attraction mutuelle entre le dipôle rigide de la première molécule et le dipôle induit de la deuxième molécule est égale d'après Debye à

$$U_{ind} = \frac{\alpha M^2}{8\pi^2\varepsilon_0^2} \frac{1}{r^6},$$

où r est la distance entre les molécules.

La liaison de Van der Waals est la plus faible — son énergie est de l'ordre de quelques kilojoules par mole. Cela explique pourquoi toutes les structures fondées sur cette liaison sont peu stables, volatiles et ont des points de fusion bas. Au point de vue électrique elles sont des diélectriques. De nombreux cristaux de ce type sont transparents aux rayonnements électromagnétiques jusqu'à l'ultraviolet lointain.

Avant de clore ce paragraphe, il convient de souligner que dans les cristaux réels chacune des liaisons qui viennent d'être considérées se rencontre rarement dans sa pureté. En pratique on a toujours une superposition de deux (ou plusieurs) types de liaisons. L'une d'elles est prédominante et détermine la structure et les propriétés du corps.

CHAPITRE 2

IMPERFECTIONS DANS LES CRISTAUX

A la différence des cristaux parfaits ayant une structure régulière rigoureusement périodique, les cristaux réels présentent diverses imperfections qui ont une forte influence pratiquement sur toutes leurs propriétés.
Parmi les défauts de structure cristalline il convient de distinguer les défauts *ponctuels*, les défauts *linéaires*, les défauts *bidimensionnels* et les défauts *tridimensionnels*. Etudions sommairement la nature de ces défauts.

§ 2.1. Défauts ponctuels

Défauts de Frenkel. La répartition de l'énergie entre les atomes d'un solide est inégale de même que la répartition entre les molécules d'un gaz et d'un liquide. A toute température il existe dans

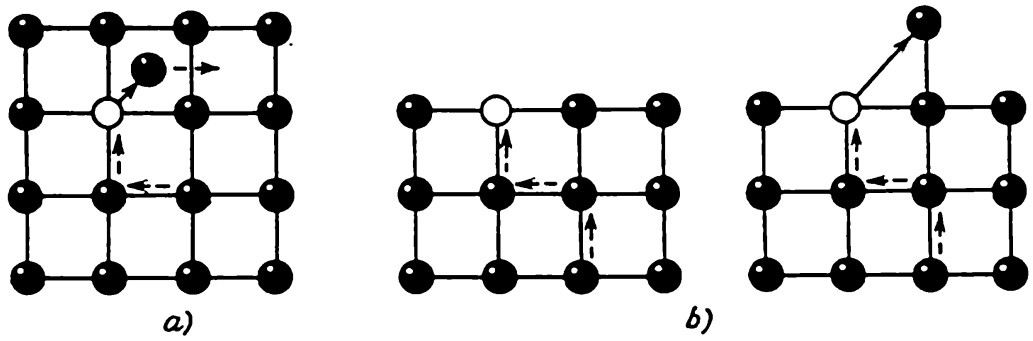

Fig. 2.1. Défauts de Frenkel (*a*) et de Schottky (*b*):
Les défauts de Frenkel se forment par suite d'une « évaporation intérieure » des atomes situés aux nœuds du réseau avec formation de vacances et par suite de leur « condensation » dans les interstices avec formation d'atomes disloqués. Aussi, se manifestent-ils toujours par paires. Les défauts de Schottky se forment par suite d'une évaporation totale ou partielle des atomes à partir de la surface avec formation de vacances seulement.

le cristal des atomes dont l'énergie est plusieurs fois plus grande ou plus petite que l'énergie moyenne déterminée par la température du cristal. Les atomes qui possèdent, à un instant donné, une énergie suffisamment haute peuvent non seulement s'éloigner à des distances considérables de la position d'équilibre mais aussi franchir la barrière de potentiel créée par des atomes voisins et passer dans un nouvel entourage, dans une nouvelle maille élémentaire. Tout se passe comme si de tels atomes avaient la possibilité de « s'évaporer » de leurs nœuds et de « se condenser » dans les interstices (fig. 2.1, *a*).

Ce processus s'accompagne d'apparition d'un nœud vacant (d'une *vacance*) et d'un atome dans l'interstice (d'un *atome disloqué*). Les défauts de ce type du réseau cristallin sont appelés *défauts de Frenkel*.

Défauts de Schottky. En plus de l'évaporation interne il est possible que les atomes s'évaporent en partie ou en totalité sur la surface du cristal. Dans l'évaporation totale l'atome quitte la surface du cristal et se résout en vapeur tandis que dans l'évaporation partielle l'atome passe de la surface dans une position au-dessus de la surface (fig. 2.1, *b*). Dans les deux cas une vacance se forme dans la couche superficielle du cristal. La substitution de la vacance par un atome situé plus profondément l'entraîne à l'intérieur du cristal. Une telle formation de vacances ne s'accompagne pas d'une introduction simultanée des atomes dans les interstices, c'est-à-dire de l'apparition des atomes disloqués. Les vacances de ce type sont appelées *défauts de Schottky*. Leur source peut être constituée aussi par toutes les imperfections possibles du cristal : les plans réticulaires incomplets, les frontières des blocs et des grains, les fissures microscopiques, etc.

La concentration d'équilibre de défauts dans le cristal dépend avant tout de la température parce que l'élévation de la température augmente le nombre d'atomes dont l'énergie devient suffisante pour vaincre la liaison avec les voisins et former un défaut. Suivant la loi de Boltzmann le nombre de tels défauts est proportionnel à $\exp[-U/(kT)]$, où U est l'énergie de formation du défaut ; T la température thermodynamique du cristal. Les calculs montrent que le nombre d'équilibre de défauts de Frenkel dans un cristal contenant N nœuds et N_1 interstices est déterminé par la relation suivante :

$$N_F = \sqrt{NN_1} \exp[-U_F/(2kT)]. \tag{2.1}$$

Le nombre d'équilibre de défauts de Schottky a pour valeur

$$N_S = N \exp[U_S/(kT)], \tag{2.2}$$

où U_F, U_S sont les énergies de formation des défauts de Frenkel et de Schottky respectivement.

Pour la formation des défauts de Frenkel l'atome passant d'un nœud dans un interstice doit non seulement rompre ses liaisons avec les atomes voisins mais encore écarter ces atomes pour s'introduire entre eux. Cela exige de dépenser une énergie considérable et ne peut se produire pratiquement que dans des cristaux constitués de deux espèces d'atomes fortement différents par leurs dimensions, par exemple dans les cristaux ioniques, les solutions solides interstitielles, etc. Dans les cristaux de structure compacte, constitués d'atomes de même type, par exemple dans les cristaux métalliques, les défauts de Frenkel ne peuvent pas pratiquement apparaître et les défauts principaux sont ceux de Schottky. Les calculs motrent que

dans un cristal de cuivre à $T = 1000$ °C la concentration de vacances $N_S/N \approx 10^{-4}$ et la concentration de défauts de Frenkel $N_F/N \approx 10^{-39}$.

En plus des vacances uniques de petites accumulations et notamment des *vacances doubles* et *triples* peuvent se former, ce qui souvent est plus avantageux au point de vue énergétique. En effet, lors de la formation d'une vacance unique l'atome qui quitte un nœud du réseau surmonte les liaisons avec z voisins (z étant le nombre de coordination du réseau). Lors de la formation d'une divacance le

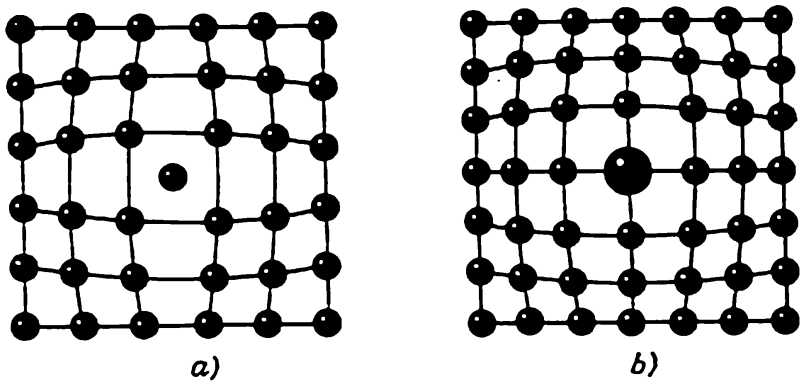

Fig. 2.2. Atomes d'impureté en position intersticielle (*a*) et en position substitutionnelle (*b*).

deuxième atome qui s'éloigne d'un nœud voisin n'a à vaincre la liaison qu'avec un plus petit nombre de voisins. Aussi l'énergie nécessaire à la formation d'une divacance doit-elle être inférieure à l'énergie nécessaire à la formation de deux vacances isolées.

Dans les composés semiconducteurs, dans $A^{III}B^{V}$ par exemple, l'apparition de la lacune d'un des constituants trouble la stœchiométrie du composé et fait apparaître un excès d'atomes de l'autre constituant. Les atomes en excès de constituant A^{III} se comportent dans le cristal comme des accepteurs et les atomes en excès de constituant B^{V} comme des donneurs.

Impuretés. Les corps solides même d'un très haut degré de pureté contiennent toujours des *impuretés*. Suivant leur nature et leur quantité les impuretés peuvent se trouver dans un cristal soit sous forme dissoute, soit sous la forme d'inclusions plus ou moins grandes. Le processus de dissolution consiste en ce que les atomes d'impureté s'introduisent dans les intervalles entre les atomes du cristal (fig. 2.2, *a*) ou se substituent à une partie de ces atomes en occupant les nœuds du réseau (fig. 2.2, *b*). Dans le premier cas, la solution solide est dite *interstitielle* et dans le second cas *substitutionnelle*. Comme la nature physique et les dimensions des atomes étrangers diffèrent de celles des atomes principaux du cristal, leur présence trouble la structure régulière du réseau cristallin.

Les impuretés peuvent avoir une influence considérable sur toutes les propriétés des solides. Elles constituent des centres de diffusion effectifs en déterminant la résistance électrique des métaux qui ne disparaît pas à 0 K (si le métal ne passe pas à l'état supraconducteur). Dans les cristaux semiconducteurs les impuretés produisent de nouveaux niveaux énergétiques et donnent lieu à une *conductibilité d'impureté*.

Les méthodes de purification modernes permettent d'obtenir des matériaux avec une teneur d'impuretés jusqu'à 10^{-9}% (10^{11} atomes d'impureté par 1 cm^3). Pourtant les exigences concernant l'élévation du degré de pureté augmentent toujours et dans l'avenir le

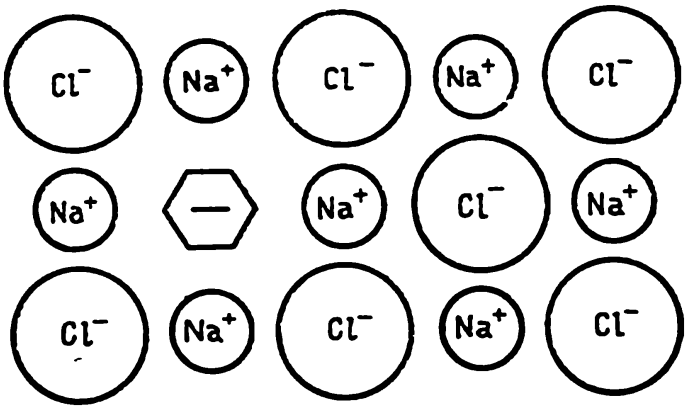

Fig. 2.3. Schéma du modèle d'un centre *F* représentant une vacance d'anion ayant capté un électron (emprunté au livre de C. Kittel, *Introduction à la physique du solide*).

plus proche on demandera des matériaux avec une teneur d'impureté de 10^{-10} à 10^{-12}%. Cela concerne avant tout les matériaux utilisés dans les installations de synthèse thermonucléaire, les matériaux pour la microélectronique et autres branches de la technique. De tels matériaux sont difficiles non seulement à obtenir mais aussi à conserver en parfait état de pureté surtout dans le cas où ils sont soumis à un traitement quelconque.

Centres de coloration. Si un cristal transparent d'halogénure, de NaCl par exemple, est chauffé dans les vapeurs de sodium, il prend une couleur jaune caractéristique ; les cristaux de KCl chauffés dans les vapeurs de potassium deviennent rouges, etc. La coloration s'explique par l'apparition dans les cristaux de défauts ponctuels particuliers qu'on appelle *centres de couleur* ou *de coloration*. Ces centres absorbent la lumière dans le domaine visible du spectre.

Les centres de coloration les plus simples sont des *centres* dits *F* qui apparaissent de la façon suivante. Lorsque NaCl est chauffé dans les vapeurs de sodium, il apparaît dans le cristal un excès (par rapport à la stœchiométrie) d'atomes de Na dont la concentration peut atteindre une valeur de 10^{16} à 10^{19} cm^{-3}. Le trouble de la stœchiométrie donne lieu à l'apparition de lacunes d'anions (fig. 2.3) et d'une

charge positive de cations non compensée autour de ces lacunes. Un tel centre chargé positivement capte un électron libre d'atomes ionisés de sodium en formant un système qui rappelle, dans une certaine mesure, un atome d'hydrogène et qui est capable d absorber la lumière lors de l'excitation et de l'émettre lors du passage à l'état normal. Pour les cristaux de sodium le maximum de la bande d'absorption correspond à l'énergie de 2,7 eV (λ = 460 nm), pour le chlorure de potassium 2,2 eV (λ = 565 nm).

Les centres *F* peuvent former des complexes constitués de deux centres *F*, de trois centres *F*, etc. De tels centres complexes sont appelés centres *M* et *R* respectivement.

Un excès d'atomes d'halogène dans le cristal a pour effet la formation de lacunes de cations qui peuvent capter des trous, en formant des centres de coloration qu'on appelle *centres V* et des complexes composés de plusieurs centres.

Les centres de coloration dans le cristal peuvent être créés non seulement par l'introduction des atomes en excès (par rapport à la stœchiométrie) d'un constituant ou d'un autre de composé mais aussi par le bombardement des cristaux par des particules accélérées — électrons, neutrons et autres et par l'irradiation des cristaux par les quanta X et gamma.

§ 2.2. Défauts linéaires

Ils sont représentés par des *dislocations coin* et *vis*.

Dislocations coins. Imaginons qu'un des plans atomiques (plan *M* de la figure 2.4, *a*) ne s'étend pas à travers tout le cristal mais s'interrompt à son intérieur. Le bord de ce plan constitue un défaut linéaire (dérangement du réseau) qu'on appelle dislocation coin, le plan *M* lui-même étant dit *extraplan*. La figure 2.4, *b* montre la disposition des atomes dans un plan perpendiculaire à une dislocation coin. Le plan atomique « superflu » *OM* est un extraplan et le point *O*, où il est coupé, représente le centre de dislocation. La rangée atomique *1* passant par le point *O* contient un atome de plus que la rangée *2* située plus bas. Ainsi la distance entre les atomes de la rangée *1* près du centre de dislocation *O* est plus petite que la distance normale de sorte que le réseau est ici serré ; quant à la distance entre les atomes de la rangée *2* près du point *O*, elle est plus grande que la distance normale de sorte qu'ici le réseau est étiré. A mesure qu'on se déplace à partir du centre de dislocation vers la droite ou vers la gauche, vers le haut ou vers le bas, la déformation du réseau diminue progressivement et, à une certaine distance de *O*, une disposition normale des atomes se rétablit dans le cristal. Dans la direction perpendiculaire au plan de la figure la dislocation passe par tout le cristal et sort sur sa surface.

Si l'extraplan se situe dans la partie supérieure du cristal, comme l'indique la figure 2.4, *b*, la dislocation est considérée comme positive

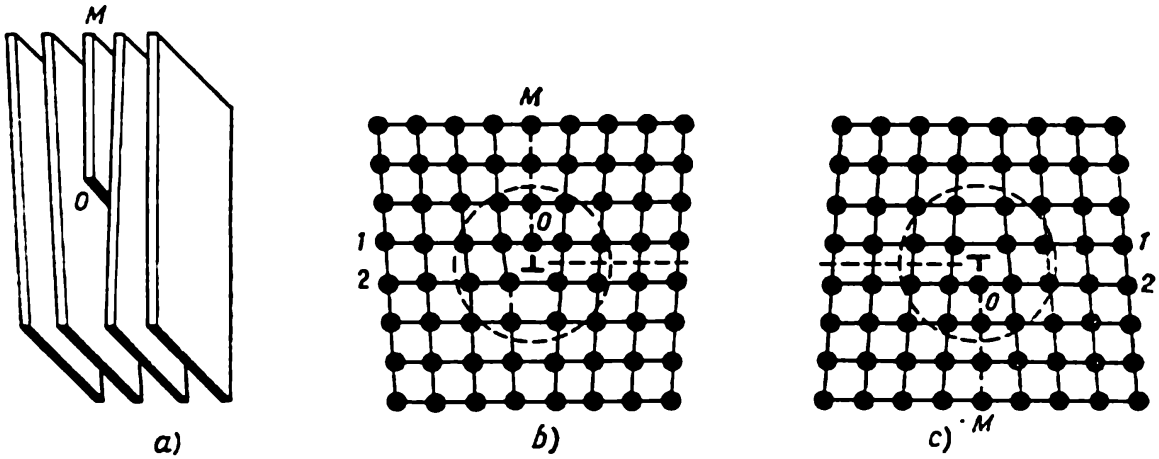

Fig. 2.4. Schéma d'une dislocation coin:
a — le bord du plan *M*, qui s'interrompt à l'intérieur du cristal, forme une dislocation coin; *b* — disposition des atomes près du centre *O* d'une dislocation positive; *c* — idem près du centre d'une dislocation négative. La région de dislocation est entourée d'un cercle en pointillé.

et marquée de signe ⊥, s'il se situe dans la partie inférieure du cristal (fig. 2.4, *c*), la dislocation est dite négative et désignée par le signe ⊤. Les dislocations de même nom, situées dans un même plan atomique, se repoussent, les dislocations de nom contraire s'attirent.

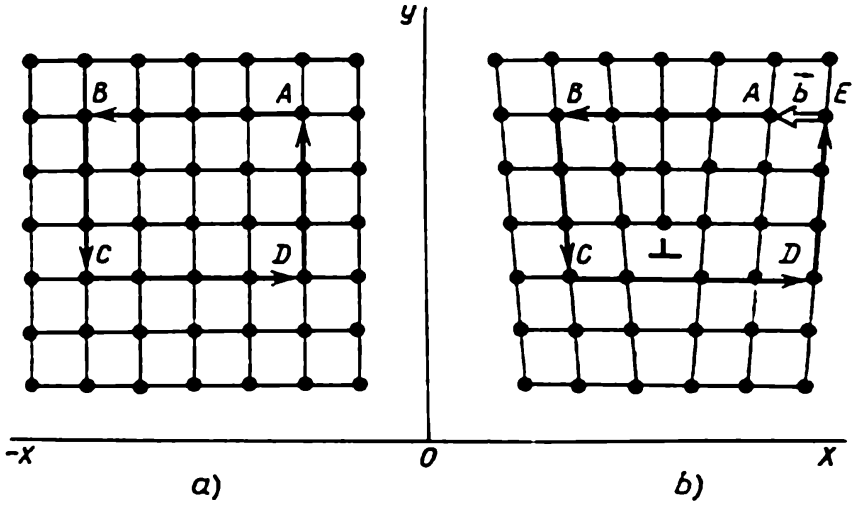

Fig. 2.5. Construction d'un contour et d'un vecteur de Burgers:
a — contour de Burgers *ABCD* pour une région de cristal ne contenant pas de dislocation coin; *b* — contour de Burgers *ABCDE* pour une région de cristal contenant une dislocation coin; **b** — vecteur de Burgers.

Quantitativement les dislocations sont décrites à l'aide de *contour* et de *vecteur de Burgers*. La figure 2.5, *a* montre la construction d'un tel contour pour une partie du cristal qui ne contient pas de

dislocation : en sortant du point A et en effectuant n déplacements atomiques dans le sens des x négatifs, puis m déplacements atomiques dans le sens des y négatifs, puis n déplacements dans le sens des x positifs et m déplacements dans le sens des y positifs, on revient au point A après avoir décrit un contour fermé $ABCDA$. En effectuant un tel parcours de la région du cristal contenant une dislocation coin (fig. 2.5, *b*), on obtient un contour non fermé $ABCDE$. Le vecteur **b** nécessaire pour fermer ce contour est appelé vecteur de Burgers, et le contour lui-même est dit contour de Burgers. Les dislocations dont le vecteur de Burgers est égal au vecteur de translation ou est son multiple entier sont dites *totales*; les dislocations dont le vecteur de Burgers est plus petit que le vecteur de translation sont dites *partielles*.

Comme le montre la figure 2.5, *b*, dans le cas de la dislocation coin le vecteur de Burgers est perpendiculaire à la ligne de dislocation.

Dislocations vis. Imaginons que sous l'action des contraintes tangentielles τ il s'est produit suivant le plan P un décalage d'une partie du cristal par rapport à l'autre, comme il est montré sur la figure 2.6, *a*. L'examen de la figure montre que dans la plus grande partie du cristal le décalage s'est effectué d'une distance atomique de sorte que loin de la ligne OO' jusqu'à laquelle s'est propagé le décalage, est conservée une disposition régulière des atomes propre à un cristal non perturbé. Au voisinage de la ligne OO', les atomes des plans atomiques situés à droite et à gauche du plan de décalage P sont décalés les uns par rapport aux autres. Ce décalage provoque une déformation locale du réseau qu'on appelle dislocation vis. La ligne OO' est appelée *axe* ou *ligne de dislocation*.

La figure 2.6, *a* montre un contour de Burgers. On voit sur cette figure que pour fermer ce contour il est nécessaire d'introduire un vecteur supplémentaire $\mathbf{b} = \overrightarrow{FA}$ parallèle à la ligne de dislocation. Ce vecteur est celui de Burgers. A la différence de la dislocation coin où le vecteur **b** est perpendiculaire à la ligne de dislocation, dans la dislocation vis le vecteur **b** est parallèle à cette ligne.

On distingue les dislocations vis à droite et à gauche (fig. 2.6, *b*). Leur déplacement dans des sens opposés conduit au décalage d'un même signe.

Puisque le réseau cristallin dans la région de dislocation est déformé, dans cette région prennent naissance des contraintes élastiques σ_{disl} et est emmagasinée une énergie élastique E_{disl} du réseau déformé. Pour une dislocation coin on peut les exprimer par des formules approchées suivantes :

$$\sigma_{disl} \approx \frac{Gb}{2\pi r}\ ; \tag{2.3}$$

$$E_{disl} \approx \frac{Gb^2}{2} L, \tag{2.4}$$

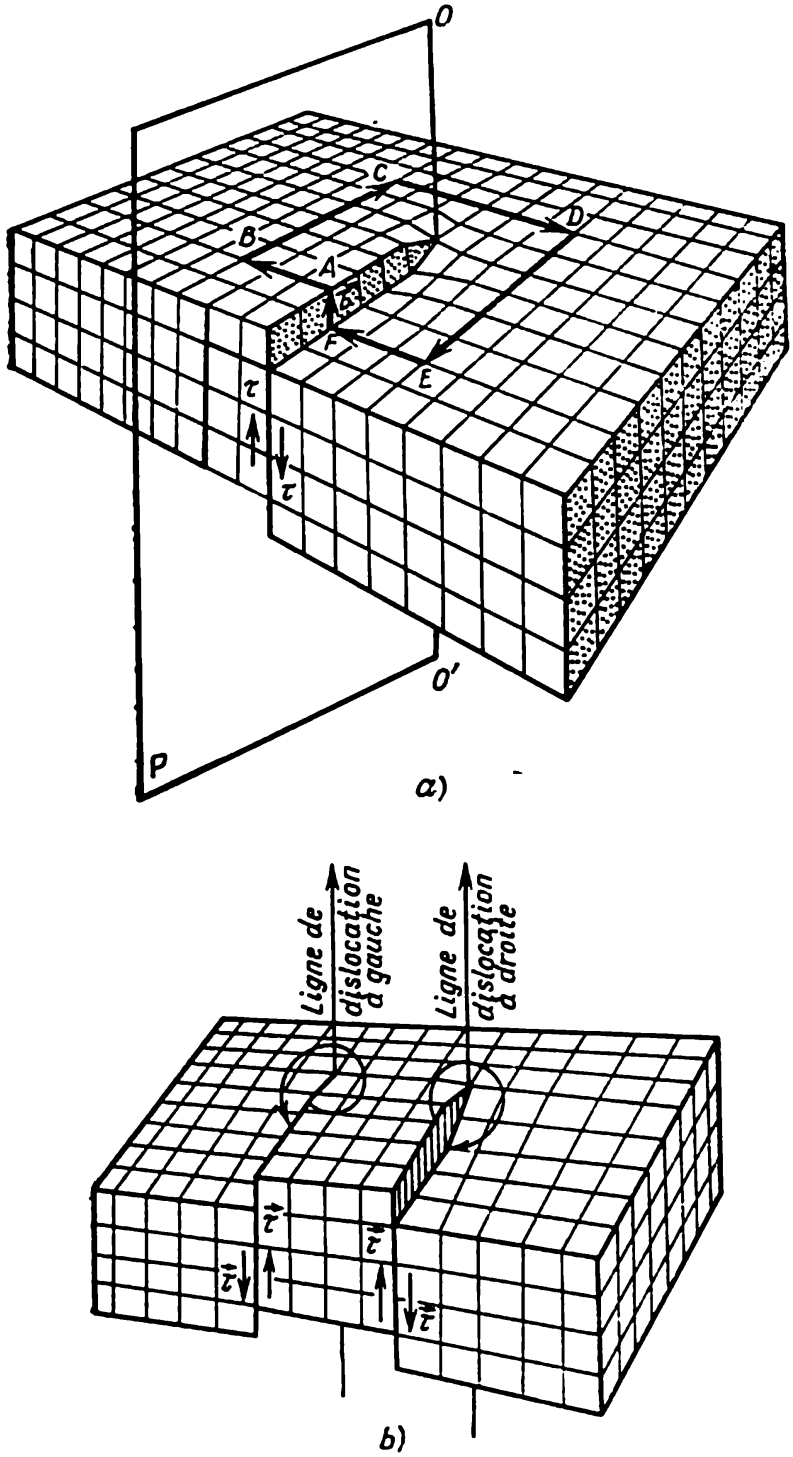

Fig. 2.6. Contour de Burgers *ABCDEF* et vecteur de Burgers **b** pour une dislocation vis (*a*) ; *b* — dislocations vis à droite et à gauche.

où G est le module de cisaillement ; **b** le module du vecteur de Burgers ; r la distance de la ligne de dislocation au point considéré ; L la longueur de la dislocation. En posant $G = 2 \cdot 10^{10}$ Pa et $b = 3 \cdot 10^{-10}$ m, on obtient $E_{disl} = 4$ eV par chaque tronçon de dislocation égal à la distance interatomique ; les contraintes élastiques à la distance

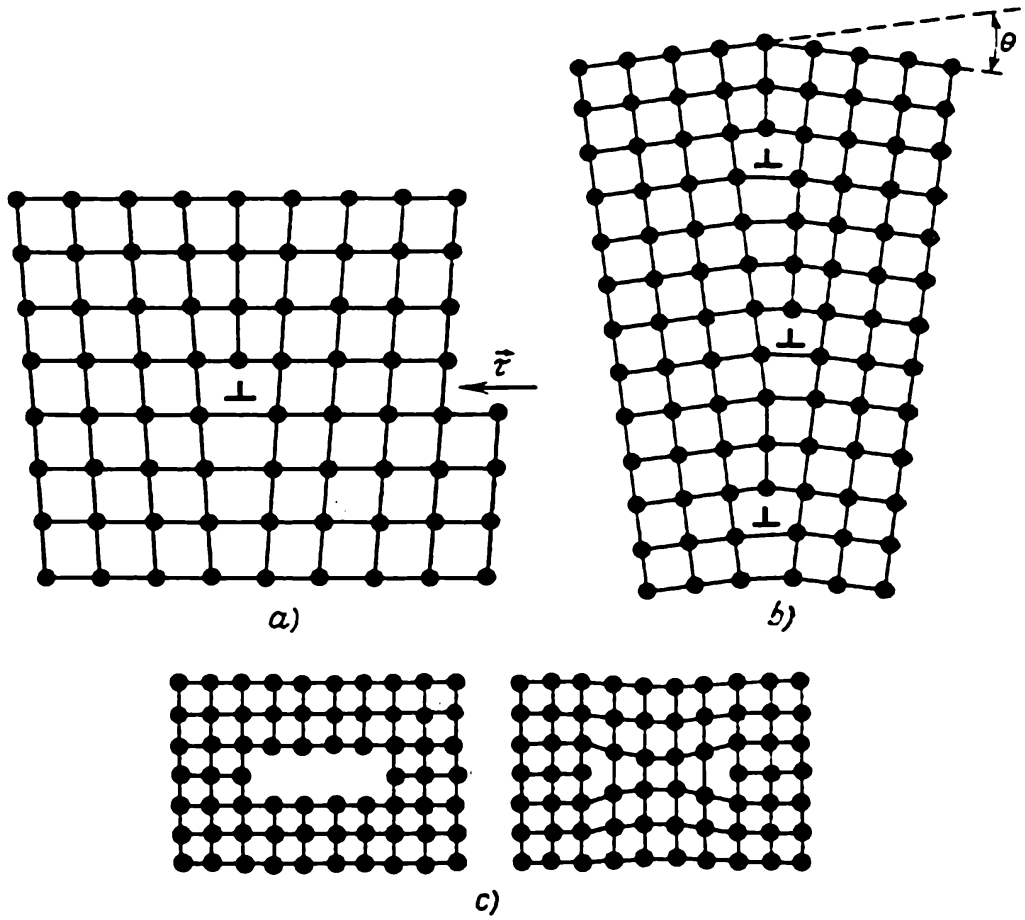

Fig. 2.7. Apparition de dislocations dans un cristal :
a — lors d'une déformation plastique (par suite d'un décalage partiel d'une partie du réseau par rapport à l'autre sous l'action des contraintes tangentielles τ) ; *b* — sur les frontières des blocs de monocristaux et de grains de polycristal (θ — l'angle fait entre les blocs) ; *c* — en cas de groupement de vacances.

$r = 100$ nm (≈ 300 distances atomiques) de la ligne de dislocation sont égales à $\sigma_{disl} \approx 10^7$ Pa (≈ 100 kgf/cm^2).

L'existence d'un champ de force autour des dislocations donne lieu à une interaction de ces dislocations avec les atomes d'impureté et d'autres défauts ponctuels. Notamment, les atomes d'impureté de petit rayon diffusent vers les régions serrées du réseau qui apparaissent près des dislocations et les atomes de grand rayon diffusent vers les régions étirées du réseau. Il en résulte que les dislocations se trouvent souvent « enveloppées » de nuages d'atomes d'impureté dits *nuages de Kottrel* qui peuvent jouer un rôle important dans la

détermination des propriétés des cristaux, dans le caractère de microdistribution des impuretés dans les cristaux, etc. Les dislocations constituent des drains pour les lacunes.

Les dislocations sont engendrées au cours de la croissance des cristaux, lors de la déformation plastique (fig. 2.7, *a*), sur les frontières des blocs de monocristaux et les frontières des polycristaux (fig. 2.7, *b*), à partir des groupements de vacances (fig. 2.7, *c*), etc. Actuellement, on a mis au point plusieurs méthodes d'observation expérimentale des dislocations : méthode du microscope électronique, méthode de décoration qui consiste à disposer sur les dislocations des atomes d'impuretés qui rendent les dislocations visibles au microscope, et autres.

La sortie des dislocations sur la surface du cristal peut être décelée par la *méthode de décapage*. Lorsque la surface du cristal est attaquée par des décapants spéciaux, ce sont les parties dans lesquelles le réseau est déformé le plus fortement qui sont gravées les premières parce que les atomes de ces parties possèdent une énergie libre excédentaire et sont donc chimiquement plus actifs. De telles parties sont justement des endroits où les dislocations sortent sur la surface et se présentent, après le décapage, sous forme de trous de décapage.

Comme caractéristique quantitative du nombre de dislocations on prend la *densité de dislocations* égale au nombre de lignes de dislocation traversant une unité de surface du cristal. Dans les cristaux de silicium et de germanium les plus perfectionnés la densité de dislocations est d'environ de 10^2 à 10^3 cm^{-2} tandis que dans des cristaux métalliques non déformés de bonne qualité elle varie de 10^5 à 10^6 cm^{-2} et dans des cristaux fortement déformés de 10^{11} à 10^{13} cm^{-2}.

§ 2.3. Mécanisme de dislocation de la déformation plastique des cristaux

Résistance au cisaillement théorique des cristaux. Le principal mécanisme de l'écoulement plastique des cristaux est la formation de cisaillement. On considérait pendant longtemps que cette formation s'effectuait par un décalage rigide d'une partie du cristal par rapport à l'autre, simultanément suivant tout le plan de glissement (fig. 2.8, *a*). Evaluons grossièrement la contrainte tangentielle nécessaire à la réalisation d'un tel décalage.

Dans un réseau non déformé, les atomes de deux plans atomiques voisins occupent des positions correspondant au minimum d'énergie potentielle. La force d'interaction entre ces atomes est nulle. Lorsque l'un des plans atomiques se déplace progressivement par rapport à l'autre, il naît des forces tangentielles qui s'opposent à ce glissement et tendent à rétablir l'équilibre troublé. En admettant pour ces forces la loi sinusoïdale de variation de τ en fonction du décalage (fig. 2.8, *b*), on peut exprimer la résistance au cisaillement de la

façon suivante :

$$\tau = A \sin (2\pi x/b), \tag{2.5}$$

où x est le décalage des atomes par rapport à la position d'équilibre ; A une constante qu'on peut déterminer à partir des considérations suivantes. Pour de petits décalages $\sin(2\pi x/b) \approx 2\pi x/b$ et donc $\tau \approx A\ (2\pi x/b)$. D'un autre côté, pour de petits décalages est valable la loi de Hooke : $\tau = G\ (x/d)$, où G est le module de cisaillement ;

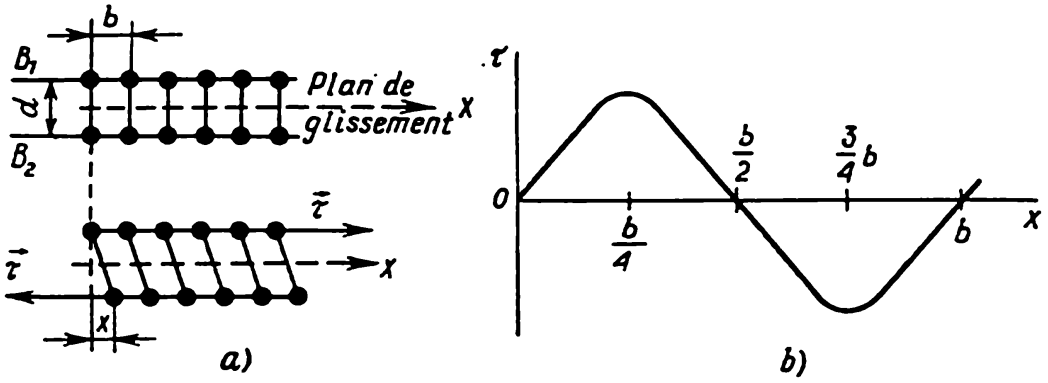

Fig. 2.8. Calcul de la résistance théorique des cristaux au cisaillement : *a* — en cas de décalage d'un plan atomique (B_1) par rapport à un autre (B_2) il y a apparition de contraintes tangentielles τ qui tendent à rétablir l'état d'équilibre troublé ; *b* — loi sinusoïdale de variation de τ avec la distance x.

d la distance entre les plans atomiques B_1 et B_2. En égalant entre eux les seconds membres de deux dernières égalités, on obtient $A = Gb/(2\pi d)$. En portant cette valeur dans (2.5) on trouve

$$\tau = \frac{Gb}{2\pi d} \sin \frac{2\pi x}{b}.$$

Pour $x = b/4$ la contrainte τ atteint sa valeur maximale $\tau_{max} = \frac{Gb}{2\pi d}$. En posant $b = d$, on obtient

$$\tau_{max} \approx \frac{G}{2\pi}. \tag{2.6}$$

De (2.6) il résulte que la résistance théorique des cristaux au cisaillement, calculée dans l'hypothèse où le décalage d'une partie du cristal par rapport à une autre se produit simultanément suivant tout le plan de glissement, doit être égale à $\approx 0{,}1\ G$, c'est-à-dire de l'ordre de 10^9 à 10^{10} Pa (10^4 à 10^5 kgf/cm^2). Une prise en compte plus rigoureuse du caractère des forces d'interaction entre les atomes ne conduit qu'à une précision peu importante de cette valeur.

Pourtant la résistance réelle des cristaux au cisaillement est, comme le montre l'expérience, inférieure de deux à quatre ordres de grandeur à cette valeur théorique, ce qui prouve que le décalage dans les cristaux se produit non par un décalage rigide des plans atomiques l'un par rapport à l'autre, mais par un mécanisme tel qu'à chaque

instant c'est seulement un nombre d'atomes relativement petit qui se déplacent. Cela a conduit au développement de la théorie de dislocation de l'écoulement plastique des cristaux.

Mécanisme de dislocation de l'écoulement plastique des cristaux. Supposons que dans un plan S de glissement possible du cristal se

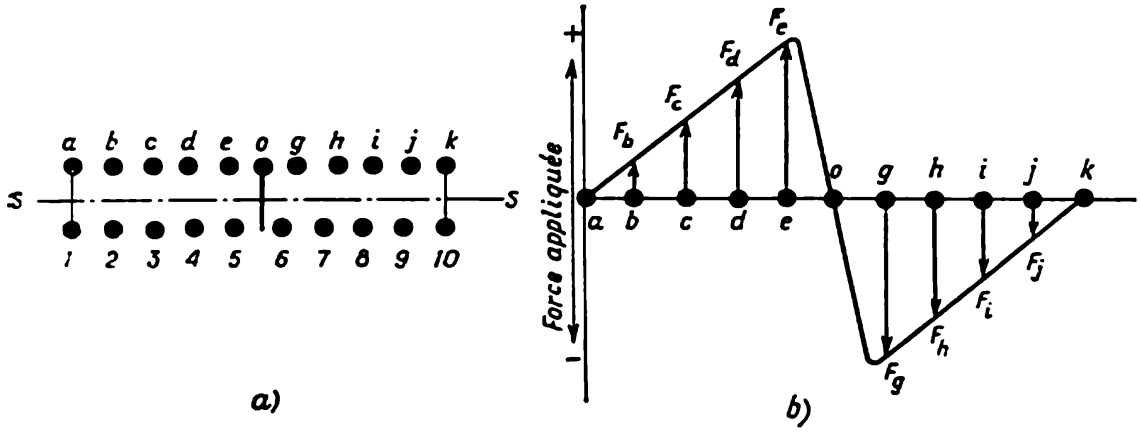

Fig. 2.9. Calcul de la force nécessaire pour le déplacement d'une dislocation: *a* — région de dislocation positive dans un cristal: *O* — centre de dislocation, *a* et *k* — frontières de la dislocation, *S* — plan de glissement ; *b* — forces nécessaires pour le déplacement d'un atome dans la région de dislocation. Les forces appliquées à des atomes équidistants du centre de dislocation sont égales en valeur absolue et de sens opposés.

trouve une dislocation positive de centre au point O et de frontières aux points a et k (fig. 2.9, *a*). A l'état d'équilibre, la force résultante agissant sur cette dislocation de la part du réseau est nulle. Cela est facile à comprendre en examinant le modèle à galets représenté par

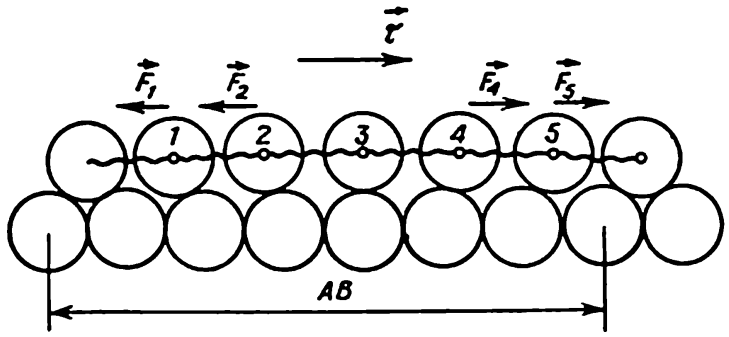

Fig. 2.10. Modèle de galets de la dislocation linéaire. Les forces appliquées aux galets *1* et *5*, *2* et *4* sont égales en valeur absolue et de sens opposés.

la figure 2.10. Dans la rangée supérieure de galets qui se disposent normalement dans les creux de la rangée inférieure, on a effectué un « dérangement de structure » : sur le tronçon AB qui était précédemment occupé par *6* galets on n'a placé que *5* galets. Un tel dérangement fait naître des forces F_1, F_2, F_4, F_5 qui tendent à faire revenir les galets *1*, *2*, *4*, *5* dans l'état d'équilibre stable. Les forces appli-

quées aux galets *1*, *5* et *2*, *4* sont égales en valeur absolue et dirigées dans des sens opposés. C'est pourquoi, si les galets de la rangée supérieure sont reliés l'un à l'autre par des ressorts élastiques, jouant le rôle de liaisons, les forces $\mathbf{F}_1$ et $\mathbf{F}_5$, $\mathbf{F}_2$ et $\mathbf{F}_4$ font équilibre les unes aux autres et le système sera en équilibre.

Une situation analogue se présente aussi dans le cas de la dislocation représentée schématiquement sur la figure 2.9, *a*: les forces qui s'exercent sur les atomes de la rangée supérieure disposés symétriquement par rapport au centre de dislocation *O* (fig. 2.9, *b*) sont égales en grandeur et de signe contraire (les forces $|F_b| = |F_j|$, $|F_c| = |F_i|$, $|F_d| = |F_h|$, $|F_e| = |F_g|$). Aussi la résultante de ces forces est-elle nulle et la dislocation est en équilibre. Toutefois, si la dislocation subit un petit déplacement dans le plan de glissement, la symétrie de disposition des atomes par rapport au centre de dislocation sera troublée, ce qui fera naître une force qui s'opposera au mouvement de la dislocation. Comme le montre la figure 2.10, cette force ne peut pas être grande parce que le déplacement des galets *1*, *2* dans de nouvelles positions d'équilibre se produit pour une large part sous l'influence des forces agissant de la part des galets *4*, *5* qui tendent eux aussi à prendre une position d'équilibre stable. Les calculs montrent que la contrainte tangentielle nécessaire pour déplacer la dislocation est égale à $\tau \approx 10^{-4}$ G. En ordre de grandeur elle coïncide avec des valeurs expérimentales de la contrainte de cisaillement qui provoque une déformation plastique dans des monocristaux métalliques de bonne qualité.

§ 2.4. Défauts bidimensionnels

Surfaces libres. Tout atome se trouvant à l'intérieur du cristal se caractérise par la symétrie des forces qui s'exercent sur lui de la part des atomes voisins. Une telle symétrie est troublée pour des atomes disposés sur la surface parce que ses voisins ne se disposent que d'un côté de cette surface. Il en résulte une déformation de l'empilement des atomes superficiels — une variation de la distance interatomique et leur décalage par rapport aux positions occupées par les atomes dans les couches atomiques à l'intérieur du cristal, parallèles à la surface. En outre, les liaisons caractéristiques du volume du cristal peuvent être troublées et remplacées par des liaisons superficielles spécifiques. C'est ainsi que sur la surface (100) des cristaux de silicium, de germanium et de diamant il se produit la rupture de deux liaisons sp^3 hybrides avec formation de deux électrons libres par atome. La mise en commun de ces électrons a pour effet la réunion supplémentaire deux à deux des atomes superficiels et leur rapprochement l'un de l'autre.

Frontières. Les cristaux réels possèdent souvent une structure en mozaïque: ils sont constitués de blocs réguliers mais leur dispo-

sition n'est qu'approximativement parallèle. Les dimensions des blocs varient approximativement de 10^{-6} à 10^{-4} cm et les angles qu'ils font entre eux sont compris entre quelques secondes et quelques minutes. Comme les réseaux dans les blocs mis en contact sont orientés d'une façon légèrement différente, on voit apparaître dans les régions de passage la formation de dislocations (fig. 2.7, *b*) dont la densité est d'autant plus grande que l'angle de désorientation des blocs est plus élevé. Des angles de désorientation encore plus élevés

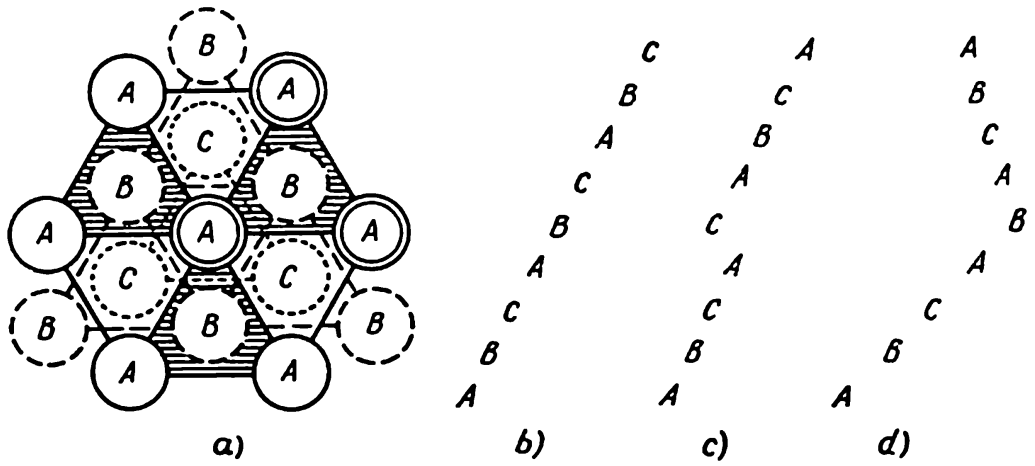

Fig. 2.11. Défauts d'assemblage:
a, *b* — disposition normale des plans atomiques [111] dans un réseau cubique à faces centrées; *c* — défaut d'assemblage du type soustraction; *d* — défaut d'assemblage conduisant au maclage.

(jusqu'à quelques dizaines de degrés) prennent naissance près des frontières des grains de polycristaux où la densité de dislocations est encore plus haute.

Les frontières de blocs et de grains sont des porteurs d'énergie libre excédentaire qui provoque une vitesse augmentée de déroulement des réactions chimiques, des transformations polymorphes, de la diffusion, etc. Elles constituent également des centres effectifs de diffusion des porteurs de charges libres déterminant une part considérable de la résistance électrique des corps.

Défauts d'assemblage. On appelle défauts d'assemblage tout écart par rapport à la disposition normale, propre au type donné de cristaux, des plans atomiques. La figure 2.11, *a* montre à titre d'exemple la suite normale de superposition des couches atomiques dans les plans {111} d'un cristal cubique à faces centrées. Dans les alvéoles hachurés de la rangée atomique inférieure *A* (cercles en traits pleins) sont disposés les atomes de la couche suivante *B* (cercles en traits interrompus); les alvéoles non hachurés restent libres. Les atomes de la troisième couche *C* (cercles en pointillé) se disposent dans les alvéoles de la couche *B* situés au-dessus des alvéoles non hachurés de la couche inférieure *A*. La quatrième couche située au-

dessus de la couche *C* (cercles en double ligne continue) répète la couche *A* et ainsi de suite.

Ainsi, la disposition normale des couches atomiques peut être représentée par un schéma montré sur la figure 2.11, *b*. Toutefois, lors de la croissance des cristaux ou en cas de leur déformation, une telle disposition des couches atomiques propres aux cristaux parfaits peut être troublée. Le cas le plus simple d'un tel dérangement est représenté sur la figure 2.11, *c* : la couche *A* est suivie de la couche *C* au lieu de la couche *B*. Un tel défaut d'assemblage est appelé *défaut du type soustraction*. La figure 2.11, *d* montre un défaut d'assemblage qui conduit à la formation de macles. D'autres types de défauts d'assemblage sont également possibles.

Défauts tridimensionnels. Ce sont des pores de divers types, des microformations, des inclusions, etc.

CHAPITRE 3

VIBRATIONS THERMIQUES DES ATOMES DE CRISTAL

Les vibrations thermiques des atomes de corps solides déterminent les propriétés importantes de ces corps, telles que la dilatation thermique, la chaleur spécifique, la conductibilité thermique. Mais pour l'électronique des solides c'est l'influence des vibrations thermiques sur les propriétés électriques des solides qui importe le plus. C'est à l'étude de toutes ces questions qu'est consacré le chapitre actuel.

§ 3.1. Vibrations et ondes dans un réseau unidimensionnel constitué d'atomes identiques

Les atomes des corps solides effectuent des vibrations thermiques autour des positions d'équilibre. Vu une forte interaction entre elles-mêmes, le caractère de ces vibrations est bien complexe et sa description précise présente des difficultés énormes. C'est pourquoi, pour résoudre ce problème, on a recours à des méthodes approchées et à toute sorte de simplifications. Tout d'abord on considère que si la distance a entre les atomes voisins varie d'une certaine quantité u par rapport à la distance d'équilibre, il se crée entre ces atomes une force F proportionnelle à u et dirigée dans le sens opposé à u:

$$F = -\beta u. \tag{3.1}$$

Une telle approximation est dite harmonique. Elle est justifiée pour de petites variations des distances interatomiques. Le cristal se réduit dans ce cas à un ensemble de particules reliées entre elles par des ressorts parfaits (fig. 3.1). Dans une telle structure les atomes ne peuvent pas se déplacer indépendamment l'un de l'autre — les vibrations de tout atome sont transmises aux atomes voisins et se propagent à travers le cristal sous la forme d'une onde élastique.

Par raison de simplification, considérons une onde élastique longitudinale plane se propageant suivant la direction [100] dans un cristal cubique. Dans ce cas les atomes dans chacun des plans perpendiculaires à la direction [100], c'est-à-dire dans tout plan (100), se déplacent en phase. Pour un certain plan s l'équation du mouvement est de la forme

$$N_{pl} M\ddot{u}_s = F_{pl\,s} = N_{pl}F_s, \tag{3.2}$$

où N_{pl} est le nombre d'atomes dans le plan ; M la masse de l'atome ; $F_{pl\,s}$ la force agissant sur le plan s de la part des plans voisins ;

F_s la force s'exerçant sur un atome de la part de ses voisins. En simplifiant les deux membres de cette équation par N_{pl}, on obtient l'équation du mouvement de l'atome s dans la chaîne ou ligne d'atomes (fig. 3.2, *a*, *b*).

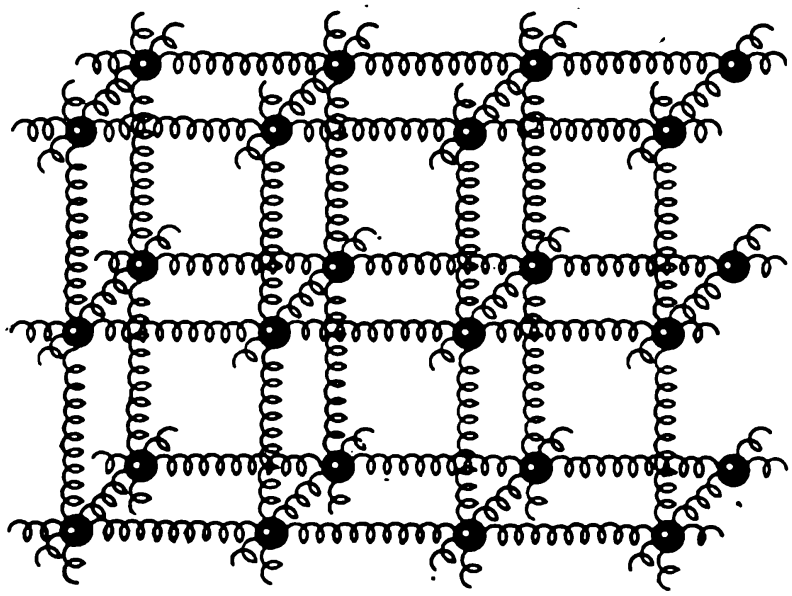

Fig. 3.1. A l'approximation harmonique le cristal peut être considéré comme un ensemble de particules disposées de façon ordonnée et reliées entre elles par des « ressorts » parfaits.

La force F_s dans une telle chaîne est déterminée par la différence de déplacements depuis les positions d'équilibre de l'atome s et des atomes $s-1$ et $s+1$:

$$F_s = M\ddot{u}_s = \beta\,(u_{s+1} - u_s) - \beta\,(u_s - u_{s-1}) = \beta\,(u_{s+1} - 2u_s + u_{s-1}). \qquad (3.3)$$

Cherchons la solution de cette équation sous la forme d'une onde progressive :

$$u_s = Ae^{iksa}e^{-i\omega t}, \qquad (3.4)$$

où a est la distance d'équilibre entre les atomes dans la chaîne ; $k = 2\pi/\lambda$ le module du vecteur d'onde ; ω la fréquence angulaire de vibrations des atomes ; A l'amplitude de vibrations. En portant (3.4) dans (3.3), on obtient

$$-\omega^2 MAe^{iksa}e^{-i\omega t} = \beta A\,(e^{ika(s+1)} - 2e^{ikas} + e^{ika(s-1)})\,e^{-i\omega t}. \qquad (3.5)$$

Après la simplification par $Ae^{ikas}e^{-i\omega t}$ on a

$$M\omega^2 = -\beta\,(e^{ika} + e^{-ika} - 2). \qquad (3.6)$$

On en déduit le lien entre la fréquence de vibrations et le vecteur d'onde, c'est-à-dire la *relation de dispersion*

$$\omega^2 = \frac{2\beta}{M}(1-\cos ka) = \frac{4\beta}{M}\sin^2\frac{ka}{2}, \tag{3.7}$$

ou

$$\omega = \sqrt{\frac{4\beta}{M}}\left|\sin\frac{ka}{2}\right|. \tag{3.8}$$

La courbe de variation de ω en fonction de k est représentée sur la figure 3.2, *c*.

Première zone de Brillouin. En réalité il est inutile de prendre une gamme de variation de k plus grande que $2\pi/a$ (par exemple

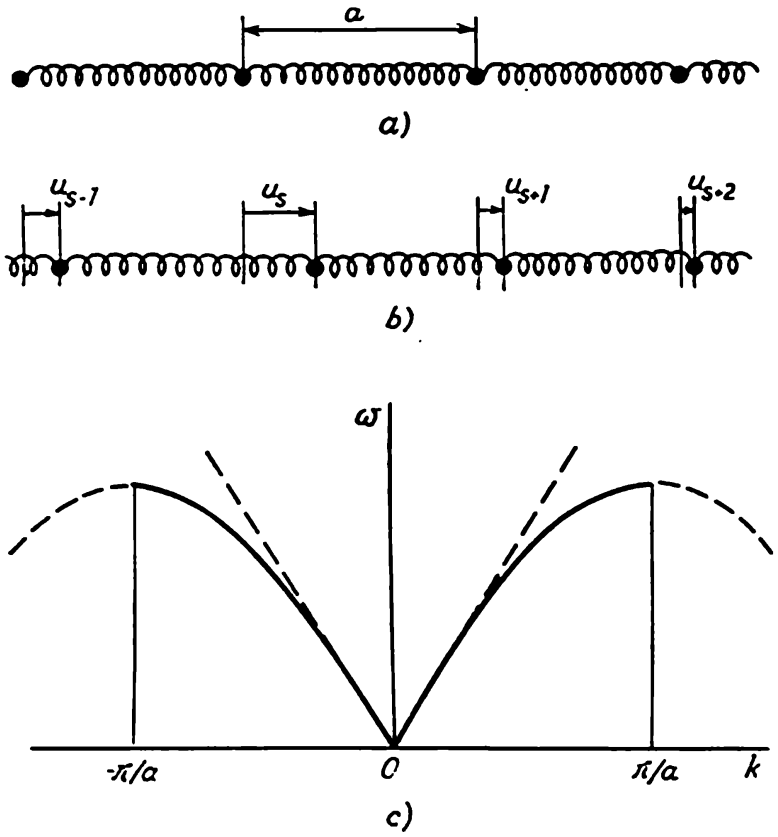

Fig. 3.2. Vibrations d'une chaîne linéaire d'atomes:
a — position d'équilibre des atomes; *b* — déplacement des atomes dans une onde longitudinale; *c* — courbe de dispersion pour une chaîne linéaire d'atomes identiques.

plus grande que $-\pi/a$ à $+\pi/a$) parce que toutes les ondes possibles physiquement différentes dans la chaîne d'atomes sont comprises dans cette collection de vecteurs d'onde. C'est à ce domaine de valeurs de k qu'on donne le nom de *première zone de Brillouin*.

Expliquons ceci sur un exemple lorsque les vibrations des atomes voisins sont en opposition de phase. Comme l'indique la figu-

re 3.3, un tel mouvement peut être décrit à l'aide d'une onde de paramètres $\lambda = 2a$, $k = \pi/a$ (ligne continue) ou d'une onde de paramètres $\lambda = \frac{2}{3}a$, $k = 3\pi/a$ (ligne en traits interrompus). Les valeurs de k diffèrent ici l'une de l'autre de $2\pi/a$, c'est-à-dire du paramètre G du réseau réciproque. On peut aussi montrer exactement de même que le même mouvement des atomes peut se décrire par une

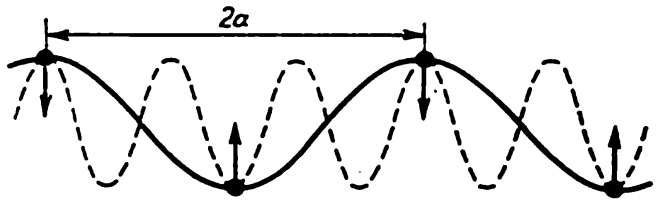

Fig. 3.3. Représentation d'une même vibration des atomes par une onde de longueur $2a$ (ligne en traits pleins) et de longueur $^2/_3a$ (ligne en traits interrompus).

onde de vecteur d'onde k qui diffère de $\pm (2\pi/a)\, n$, où n est un entier quelconque. En effet, si l'on substitue à k dans (3.4) la valeur $k \pm (2\pi/a)n$, on obtient

$$u_s = Ae^{iksa}e^{i2\pi sn}e^{-i\omega t} \equiv Ae^{iksa}e^{-i\omega t}. \qquad (3.9)$$

Conditions aux limites périodiques. La condition (3.3) n'est valable pour tout atome que dans le cas où on considère un cristal illimité. Pour ce qui est des cristaux de dimensions limitées, il est nécessaire de donner le comportement des atomes sur les surfaces, c'est-à-dire d'introduire les conditions aux limites. L'introduction des conditions aux limites peut modifier la forme des solutions obtenues. Par exemple, dans des chaînes d'atomes à extrémités libres les ondes progressives de la forme (3.4) ne peuvent pas se propager, au lieu de ces ondes il existe des *ondes stationnaires*, c'est-à-dire des combinaisons des ondes progressives de même vecteur d'onde allant à la rencontre l'une de l'autre. En outre le spectre des valeurs possibles de k varie lui aussi — au lieu d'être continu il devient discret. La collection de solutions obtenues est très grande — leur nombre est égal au nombre de degrés de liberté de tous les atomes du cristal. Sous l'effet de la chaleur il se produit dans le cristal l'excitation simultanée d'un nombre énorme de vibrations de divers types qui diffèrent l'une de l'autre par les fréquences, les directions de propagation et les polarisations. C'est précisément la participation de chacun des atomes à toutes ces vibrations à la fois qui détermine leur mouvement totalement désordonné.

Si l'on fait varier la forme des conditions aux limites, le caractère de mouvement des atomes proches de la surface changera, mais pour la masse principale d'atomes le mouvement restera aussi désordonné qu'auparavant. Ceci étant, les valeurs moyennes caractéri-

sant l'agitation thermique des atomes resteront elles aussi inchangées. Cette circonstance permet de choisir les conditions aux limites sans se préoccuper de leur réalisabilité réelle, en cherchant seulement à simplifier les calculs mathématiques. En physique du solide, on utilise généralement des conditions aux limites périodiques. Par exemple, au lieu d'une chaîne limitée de longueur L on considère une chaîne infinie divisée en tronçons de longueur L dont chacun est le siège de vibrations qui se reproduisent exactement à chaque instant. Ainsi, les déplacements des atomes situés au point x et au point $x + L$ doivent coïncider l'un avec l'autre :

$$u(x) = u(x+L), \quad \text{c'est-à-dire } e^{ikx} = e^{ik(x+L)}. \tag{3.10}$$

De (3.10) il résulte que

$$e^{ikL} = 1, \quad \text{c'est-à-dire que } k = 0;\ \pm 2\pi/L;\ \pm 4\pi/L. \tag{3.11}$$

§ 3.2. Vibrations optiques

Considérons maintenant une chaîne constituée d'atomes de deux espèces qui alternent régulièrement l'un avec l'autre (fig. 3.4, *a*). Notons M la masse des atomes plus lourds et m celle des atomes plus légers. Dans une telle chaîne deux types de vibrations montrées sur la figure 3.4, *b*, *c* peuvent prendre naissance. Les vibrations représentées sur la figure 3.4, *b* ne diffèrent pas des vibrations d'une chaîne homogène : dans le cas des ondes longues les atomes voisins vibrent pratiquement en phase et $\omega_{ac} = 0$ pour $k = 0$. De telles vibrations sont dites *acoustiques* parce qu'elles contiennent tout le spectre de vibrations sonores de la chaîne. Elles jouent un rôle principal dans la détermination des propriétés thermiques des cristaux — de la chaleur spécifique, de la conductibilité thermique, de la dilatation thermique, etc.

Dans le cas des vibrations représentées sur la figure 3.4, *c* les atomes voisins vibrent pratiquement en opposition de phase. Ces vibrations peuvent être considérées comme des vibrations, l'un par rapport à l'autre, de deux sous-réseaux constitués d'atomes homogènes qui s'emboîtent l'un dans l'autre. On les appelle vibrations *optiques* parce que leurs fréquences correspondent aux fréquences des ondes optiques infrarouges ; elles jouent un rôle prépondérant dans le processus d'interaction de la lumière avec le cristal.

La figure 3.4, *d* montre les courbes de dispersion des vibrations normales, acoustiques et optiques d'une chaîne constituée de deux espèces d'atomes. Si pour les vibrations acoustiques la fréquence augmente lorsque le vecteur d'onde augmente et atteint sa valeur maximale pour $k_{max} = \pi/a$, pour les vibrations optiques ω_{max} a lieu pour $k = 0$; lorsque k augmente la fréquence de ces vibrations diminue et devient minimale pour $k_{max} = \pi/a$.

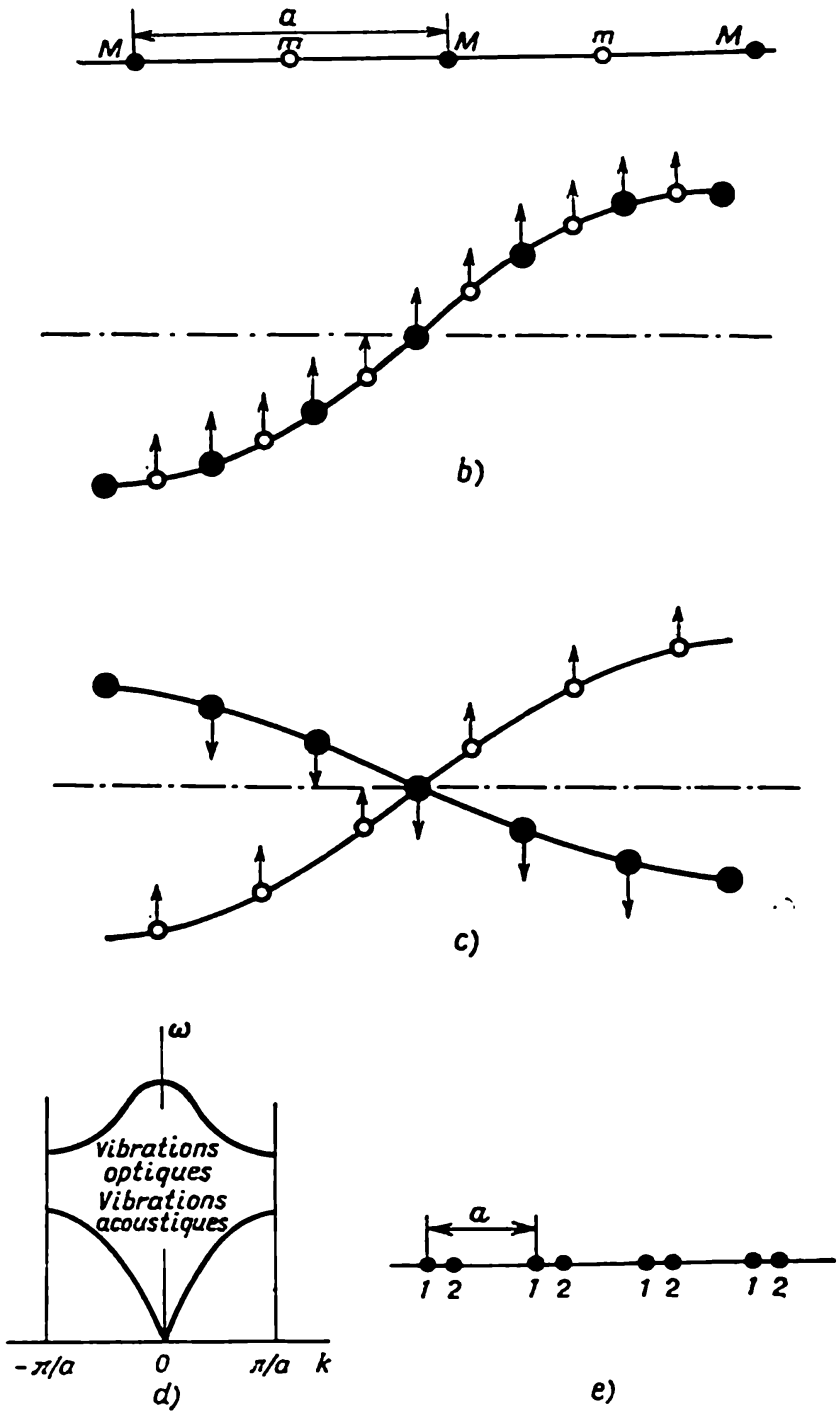

Fig. 3.4. Vibrations d'une chaîne linéaire constituée d'atomes de deux espèces: *a* — chaîne linéaire d'atomes de deux espèces; *b* — vibrations acoustiques de la chaîne; *c* — vibrations optiques de la chaîne; *d* — courbes de dispersion pour les vibrations acoustiques et optiques de la chaîne; *e* — chaîne d'atomes identiques à base pouvant devenir le siège des vibrations acoustiques et optiques.

Les vibrations optiques prennent naissance non seulement dans une chaîne composée par des atomes de nature différente, mais aussi dans le cas où la chaîne est constituée par deux (ou un plus grand nombre) de chaînes simples formées par des atomes identiques et emboîtées l'une dans l'autre comme l'indique la figure 3.4, *e*. La maille élémentaire d'une telle chaîne contient deux atomes. Les vibrations optiques ont leur origine dans les vibrations d'un sous-réseau par rapport à l'autre.

§ 3.3. Vibrations d'un réseau cristallin tridimensionnel

Considérons les vibrations dans un cristal de forme cubique, de dimensions $L \times L \times L$ dont les faces sont parallèles aux faces de la maille élémentaire. Introduisons des conditions aux limites périodiques suivant chacun des axes x, y et z. Ces conditions signifient maintenant que l'espace est divisé en cubes de dimensions $L \times L \times \times L$ et l'image des vibrations dans le cube considéré est reproduit exactement dans chacun des autres cubes. Le déplacement des atomes dans une onde sinusoïdale plane, se propageant dans une certaine direction dans le cristal, est donné sous la forme

$$\mathbf{u}(\mathbf{r}) = \mathbf{u}_0 e^{-i\omega t} e^{i\mathbf{kr}} = \mathbf{u}_0 e^{-i\omega t} e^{ik_x x} e^{ik_y y} e^{ik_z z}, \qquad (3.12)$$

où $\mathbf{u}_0$ est l'amplitude de déplacement des atomes dans l'onde; k_x, k_y, k_z sont les projections du vecteur d'onde $\mathbf{k}$ sur les axes de coordonnées.

En appliquant les conditions aux limites périodiques exactement de la même façon que pour la chaîne linéaire d'atomes, on trouve les valeurs autorisées des projections k_x, k_y, k_z du vecteur d'onde $\mathbf{k}$ sur les axes de coordonnées, égales à 0, $\pm 2\pi/L$, $\pm 4\pi/L$, et ainsi de suite jusqu'aux valeurs maximales $\pm\pi/a$ (a étant la constante du réseau cristallin).

La première zone de Brillouin du cristal cubique représente maintenant un cube dont la longueur de l'arête est égale à $2\pi/a$ dans l'*espace* dit *des* $\mathbf{k}$. Sur la figure 3.5 les points indiquent les valeurs permises du vecteur d'onde $\mathbf{k}$.

Nombre total de modes. A chaque valeur permise de $\mathbf{k}$ correspond, dans l'espace des $\mathbf{k}$, un volume $(2\pi/L)^3 = 8\pi^3/V$, où V est le volume du cristal. En divisant le volume de la première zone de Brillouin $8\pi^3/a^3$ par ce volume élémentaire, on trouve le nombre total de différentes valeurs permises du vecteur d'onde $\mathbf{k}$. Il est égal à V/a^3. Comme a^3 n'est rien d'autre que le volume de la maille élémentaire du cristal, V/a^3 est égal au nombre N d'atomes dans le cristal. Pourtant, les ondes dans un cristal peuvent différer l'une de l'autre non seulement par leurs vecteurs d'onde $\mathbf{k}$ mais aussi par la direction de déplacement des atomes dans l'onde par rapport à $\mathbf{k}$, c'est-à-dire par

l'orientation du vecteur $\mathbf{u}_0$ figurant dans (3.12). Cette direction détermine la *polarisation de l'onde*. Une onde de polarisation quelconque peut être représentée toujours sous forme de trois ondes dont les polarisations sont orthogonales entre elles: une d'elles est longitudinale et les deux autres transversales. L'onde sinusoïdale ayant une de ces trois directions de polarisation et une certaine valeur

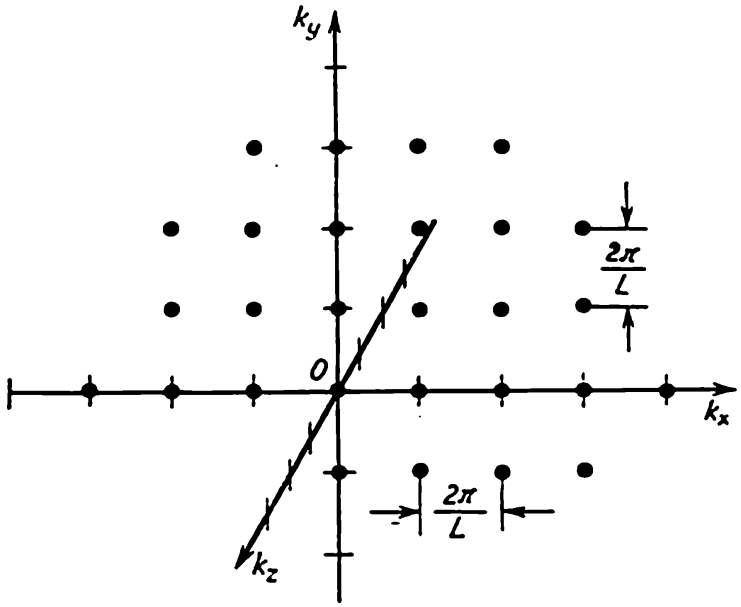

Fig. 3.5. Valeurs permises du vecteur d'onde $\mathbf{k}$ dans un cristal de dimensions $L \times L \times L$.

permise du vecteur d'onde $\mathbf{k}$ est appelée *vibration normale* ou *mode*. De ce qui précède nous pouvons conclure que le nombre total de modes différents dans le cristal est égal à $3N$.

Il est intéressant de noter que si le cristal est considéré comme un ensemble de N atomes qui se déplacent indépendamment les uns par rapport aux autres et dont chacun possède trois degrés de liberté, le nombre total de degrés de liberté d'un tel ensemble est égal à $3N$. Or, en réalité les atomes sont liés entre eux par des forces d'interaction et constituent un système vibrant unique. Néanmoins, ce système a conservé son nombre de degrés de liberté. Maintenant le cristal peut être le siège de $3N$ vibrations normales différentes dont chacune joue le rôle d'un degré de liberté du cristal.

Courbes de dispersion. Dans le cas général, la fréquence de vibrations normales dans un cristal tridimensionnel dépend non seulement de la longueur d'onde mais aussi de sa polarisation et de sa direction de propagation. Ainsi, au lieu d'une seule courbe de dispersion dans le cas de la chaîne linéaire d'atomes, il existe maintenant pour toute direction de l'onde dans le cristal trois courbes de dispersion

(fig. 3.6). Dans un cristal cubique, les courbes pour deux polarisations transversales sont confondues de sorte qu'il reste une courbe pour les ondes transversales et une autre pour les ondes longitudinales (de fréquences plus élevées). Pour des réseaux à base il se crée, en plus des vibrations acoustiques, encore des vibrations optiques correspondant aux vibrations de différents sous-réseaux l'un par rapport à l'autre.

Spectre de fréquences de vibrations du réseau cristallin. Aux valeurs permises des vecteurs d'onde **k** correspondent certaines fréquences permises des vibrations (fig. 3.7). Pour de nombreux problèmes de

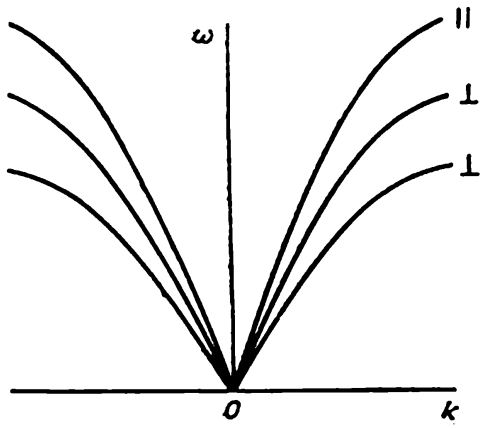

Fig. 3.6. Courbes de dispersion pour les vibrations acoustiques dans un cristal:
∥ — pour les vibrations longitudinales,
⊥ — pour les vibrations transversales.

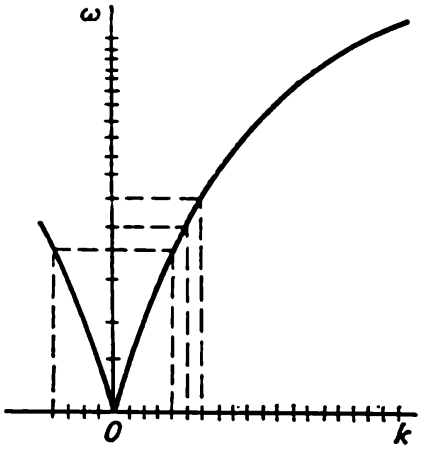

Fig. 3.7. Lien entre les valeurs permises du vecteur d'onde et le spectre de fréquences des vibrations du réseau.

la théorie du solide il importe de connaître combien de types différents de vibrations (modes) correspondent à une gamme ou à une autre de fréquences de vibrations. Etant donné que le nombre de tous les modes possibles d'un cristal de dimensions macroscopiques est astronomiquement grand, les valeurs permises se disposent sur l'échelle des fréquences d'une façon très serrée et il est donc commode d'introduire une *fonction de densité de répartition des modes suivant les fréquences* $g(\omega)$. La valeur de la grandeur $g(\omega)\,d\omega$ est égale au nombre de modes compris dans l'intervalle de ω à $\omega + d\omega$ et la valeur de $g(\omega)$ est égale au nombre de modes dans l'intervalle unitaire de fréquences au voisinage de la fréquence ω.

Le calcul de la fonction de densité de modes pour un cristal tridimensionnel est dans le cas général un problème très compliqué. Aussi effectue-t-on sa simplification considérable en introduisant les hypothèses simplificatrices suivantes: tout d'abord on admet que la fréquence de vibrations dépend uniquement du module du vecteur d'onde et ne dépend pas de la direction de propagation de l'onde. Alors, à un certain intervalle de fréquences $d\omega$ correspond un inter-

valle de nombres d'onde dk — une couche sphérique dans l'espace tridimensionnel des **k** (fig. 3.8). Quant au nombre de différentes valeurs du vecteur d'onde **k** contenues dans cette couche, on le trouve en divisant le volume de la couche $4\pi k^2 dk$ par le volume élémentaire $8\pi^3/V$. Ainsi ce nombre a pour valeur

$$dg = Vk^2 dk/(2\pi^2). \qquad (3.13)$$

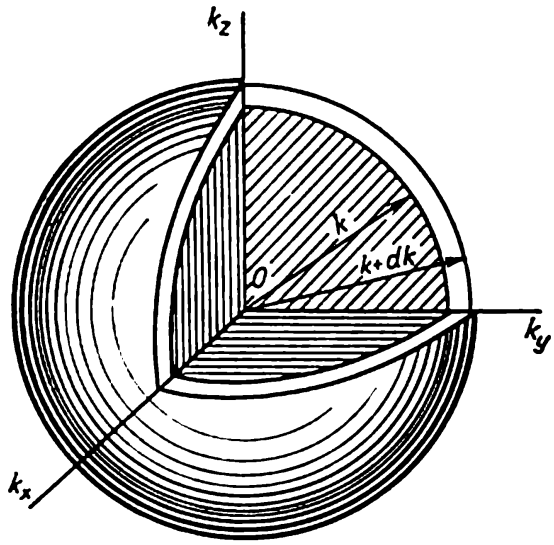

Fig. 3.8. Calcul de la fonction de densité des modes.

Maintenant, pour trouver le nombre de modes dans l'intervalle de fréquences $d\omega$ il faut remplacer k par ω suivant la loi de dispersion $\omega(k)$ et dk par $d\omega$.

Souvent on se sert des lois de dispersion les plus simples:

$$\omega = v_{\parallel} k \qquad (3.14)$$

pour les ondes longitudinales et

$$\omega = v_{\perp} k \qquad (3.15)$$

pour les deux ondes transversales. Alors un changement de variable convenable donne

$$dg_{\parallel} = \frac{V}{2\pi^2 v_{\parallel}^3}\,\omega^2\, d\omega\,; \qquad (3.16)$$

$$g_{\parallel}(\omega) = \frac{dg_{\parallel}}{d\omega} = \frac{V\omega^2}{2\pi^2 v_{\parallel}^3}\,. \qquad (3.17)$$

Pour deux polarisations transversales des ondes on a

$$g_{\perp}(\omega) = \frac{V\omega^2}{\pi^2 v_{\perp}^3}\,. \qquad (3.18)$$

La fonction complète de densité des modes est de la forme

$$g(\omega) = \frac{V\omega^2}{2\pi^2}\left(\frac{1}{v_{\parallel}^3} + \frac{2}{v_{\perp}^3}\right). \qquad (3.19)$$

L'analyse qui précède est valable tant que les surfaces sphériques dans l'espace des **k** ne commencent de sortir au-delà des faces du cube de la première zone de Brillouin. Pour éviter cette complication, Debye a proposé de considérer, au lieu de la zone cubique de Brillouin, une zone sphérique ayant le même volume et donc le même nombre total de modes $3N$. Le rayon k_D de cette sphère se détermine à partir de la relation

$$^4/_3\,\pi k_D^3 = 8\pi^3/a^3. \qquad (3.20)$$

Au vecteur d'onde k_D de Debye correspondent des fréquences de vibrations maximales pour les ondes longitudinales et pour les ondes transversales. Parfois, on pousse la simplification jusqu'à poser $v_{\parallel} = v_{\perp}$. Alors l'expression (3.19) prend la forme

$$g(\omega) = \frac{3V\omega^2}{2\pi^2 v^3} \qquad (3.21)$$

pour des fréquences allant de zéro à $\omega_D = \sqrt[3]{6\pi^2 v^3 N/V}$. Enfin, en utilisant la dernière expression pour ω_D, on peut mettre la formule (3.21) sous la forme suivante :

$$g(\omega) = \frac{9N}{\omega_D^3}\,\omega^2\,d\omega. \qquad (3.22)$$

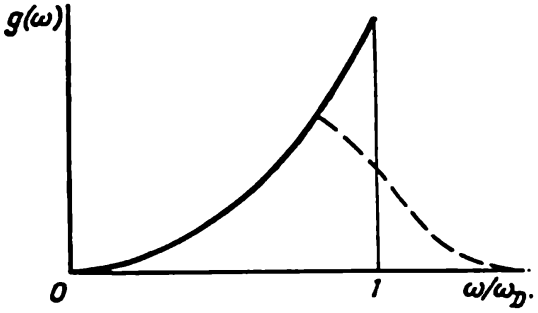

Fig. 3.9. Variation de la densité de nombre de modes en fonction de la fréquence relative de vibrations ω/ω_D à l'approximation de Debye :
courbe en traits pleins — pour la zone de Brillouin sous forme d'une sphère ; courbe en traits interrompus — pour la zone cubique de Brillouin.

La courbe représentative de la fonction de densité des modes (3.21) est représentée sur la figure 3.9 de même que la courbe traduisant la fonction de densité pour la zone cubique de Brillouin. Il y a lieu de noter que dans tous les cas la fonction de densité des modes atteint sa valeur maximale au voisinage de ω_D et, dans les cas où pour une estimation approchée il faut utiliser une certaine fréquence moyenne de vibrations du réseau, on la prend voisine de ω_D.

§ 3.4. Energie des vibrations normales. Phonons

Jusqu'ici, toutes les conclusions sur le mouvement des atomes dans les cristaux ont été faites sur la base des lois classiques du mouvement. Or, quand il s'agit de l'énergie possible de la vibration normale, une étude plus correcte, quantique, devient nécessaire. Au point de vue de la théorie classique l'énergie de la vibration normale peut prendre toute valeur et varier de façon continue par portions aussi petites que l'on veut. Quant à la mécanique quantique, elle affirme que l'énergie $E_{v.n}$ de la vibration normale de fréquence $\omega_{v.n}$ ne peut prendre qu'une suite de valeurs discrète :

$$E_{v.n} = \hbar\omega_{v.n}\,(n + {}^1/_2) \quad (n = 0,\ 1,\ 2,\ \ldots), \qquad (3.23)$$

où $\hbar$ est la constante de Planck divisée par 2π.

Une même formule est valable aussi pour l'énergie d'un oscillateur harmonique. De cette relation importante on peut tirer immédiatement deux conclusions : l'énergie minimale de la vibration normale, du mode, n'est pas nulle mais égale à ${}^1/_2\,\hbar\omega_{v.n}$; l'énergie

d'un mode ne peut varier que par des portions égales à $\hbar\omega_{v.n}$ que l'on appelle *phonons*. La première de ces conclusions peut s'énoncer aussi sous la forme suivante : *dans la nature il n'existe pas de réseaux cristallins et d'oscillateurs non vibrants.*

Les vibrations d'énergie minimale $^1/_2\hbar\omega_{v.n}$ sont appelées vibrations *de point zéro*. L'énergie de point zéro ne peut être prélevée qu'en détruisant le cristal. Lors de l'étude de nombreux processus se déroulant dans le cristal on peut négliger l'énergie de point zéro en tant que quantité invariable et utiliser au lieu de (3.23) une expression plus simple

$$E_{v.n} = n\hbar\omega_{v.n}. \tag{3.24}$$

En faisant usage de la notion de phonon, on peut, au lieu de dire que la vibration normale est excitée jusqu'au n-ième niveau vibrationnel, dire que le cristal contient n phonons d'énergie $\hbar\omega_{v.n}$ chacun. Dans ce cas le phonon joue le rôle d'une quasi-particule d'énergie bien déterminée. Etant donné que dans le réseau cristallin peuvent être excitées des vibrations normales de fréquences les plus variées, les phonons peuvent posséder des énergies différentes (v. fig. 3.7).

§ 3.5. Chaleur massique des solides

L'exactitude des déductions théoriques sur les vibrations thermiques des réseaux cristallins peut être vérifiée en confrontant les conclusions de la théorie aux expériences concernant la mesure de la chaleur massique des solides.

Théorie classique. Le plus simple est de considérer le cristal comme une assemblée d'un grand nombre N d'atomes qui vibrent, les uns indépendamment des autres, autour de leurs positions d'équilibre dans trois directions orthogonales entre elles, c'est-à-dire de considérer que chaque atome du cristal représente une petite bille fixée au moyen de trois ressorts. Alors, à chaque atome on peut faire correspondre trois oscillateurs indépendants et représenter le cristal tout entier comme un ensemble de $3N$ oscillateurs indépendants. Chacun de ces oscillateurs possède un degré de liberté du mouvement et le cristal dans son ensemble a $3N$ degrés de liberté.

Suivant la physique statistique classique, à chaque degré de liberté du mouvement du système à l'état d'équilibre thermique correspond une énergie cinétique moyenne

$$\bar{E}_{cin} = {}^1/_2\, kT, \tag{3.25}$$

où k est la constante de Boltzmann ; T la température thermodynamique *). Or la mécanique classique montre que l'énergie cinétique

*) Dans les publications consacrées à la physique du solide au lieu du terme « température thermodynamique » on utilise parfois le terme « température absolue ».

moyenne d'un oscillateur est égale à son énergie potentielle moyenne. Ainsi à chaque degré de liberté du cristal correspond une énergie totale kT et l'énergie totale des vibrations de tout le cristal est égale à

$$E_{cr} = 3NkT. \tag{3.26}$$

Il en résulte que la chaleur massique (à volume constant) du cristal a pour valeur

$$c_V = \frac{dE_{cr}}{dT} = 3Nk. \tag{3.27}$$

Cette expression ne fait intervenir que le nombre total N d'atomes du corps. La nature de ces atomes est sans importance. En d'autres termes, *les chaleurs massiques de divers corps solides contenant un même nombre d'atomes sont égales.*

En effet, il existe une *loi* empirique *de Dulong* et *Petit* (1819) qui stipule que *la chaleur massique de tout corps solide contenant 1 mole de substance simple* (pour des substances composées le nombre total d'atomes doit être égal au nombre d'Avogadro N_A) *est égale à* $3N_Ak \approx$ ≈ 25 J/(mole · K). Cette loi n'est pourtant valable que pour des températures de l'ordre de la température ordinaire et plus élevées. A basse température, la chaleur massique des solides diminue et tend vers zéro quand $T \to 0$.

Il est à noter que la loi de Dulong et Petit se déduit de la théorie classique également dans le cas où l'on tient compte de la liaison mutuelle des atomes dans le réseau cristallin. En effet, comme il a déjà été dit au paragraphe précédent, le nombre de degrés de liberté du cristal est toujours égal à $3N$. La différence consiste seulement en ce que maintenant à chaque degré de liberté correspond une vibration normale indépendante, un mode. L'énergie cinétique d'une telle vibration est la partie de l'énergie cinétique de tous les atomes du cristal qui correspond à leur moumevent dans l'onde considérée et l'énergie potentielle vaut une partie de l'énergie de déformation élastique du cristal.

Théorie quantique d'Einstein. C'est seulement dans le cadre de la théorie quantique qu'on a réussi à expliquer la chute à zéro de la chaleur massique des solides lorsque la température diminue. Einstein a montré que si l'on part de la formule de l'énergie (3.24), l'énergie moyenne de la vibration normale à l'état d'équilibre thermique a pour valeur

$$\overline{E}_{v.n} = \frac{\hbar\omega_{v.n}}{e^{\hbar\omega_{v.n}/(kT)} - 1}, \tag{3.28}$$

où $\omega_{v.n}$ est la fréquence de vibration normale. De (3.28) et (3.24) il s'ensuit que dans un tel mode il existe en moyenne $\overline{n}$ phonons

$$\overline{n} = 1/(e^{\hbar\omega_{v.n}/(kT)} - 1). \tag{3.29}$$

Si l'on suppose que tous les $3N$ modes du cristal ont une seule et même fréquence de vibrations ω, l'énergie des vibrations du cristal et sa chaleur massique ont pour expressions respectives

$$E_{cr} = 3N\overline{E}_{v.n} = \frac{3N\hbar\omega}{e^{\hbar\omega/(kT)} - 1}; \tag{3.30}$$

$$c_V = 3Nk\left(\frac{\hbar\omega}{kT}\right)^2 \frac{e^{\hbar\omega/(kT)}}{(e^{\hbar\omega/(kT)} - 1)^2}. \tag{3.31}$$

Considérons ces expressions dans deux cas extrêmes:

a) hautes températures, $kT > \hbar\omega$. Dans ce cas

$$e^{\hbar\omega/(kT)} \approx 1 + \frac{\hbar\omega}{kT};$$

$$E_{cr} \approx 3NkT; \quad c_V \approx 3Nk.$$

Ainsi aux températures élevées la théorie d'Einstein est en accord avec la loi de Dulong et Petit;

b) basses températures, $kT \ll \hbar\omega$. Dans ce cas on a suivant (3.31) $c_V \sim \exp[-\hbar\omega/(kT)]$. Cette expression tend vers zéro quand $T \to 0$. Ainsi la théorie d'Einstein décrit qualitativement correctement la variation de la chaleur massique aussi aux basses températures. Toutefois l'expérience montre que dans ce domaine des températures la chaleur massique varie suivant la loi $c_V \sim T^3$. Cette contradiction a été résolue par Debye qui a tenu compte du fait que les fréquences des vibrations normales ne sont pas égales mais occupent un large intervalle du spectre.

Théorie de Debye. Le nombre de vibrations normales contenues dans un intervalle de spectre $d\omega$ est égal à $g(\omega)\,d\omega$. En multipliant ce nombre par l'énergie moyenne de la vibration normale, on obtient l'énergie totale des vibrations normales comprises dans l'intervalle $d\omega$:

$$dE_{cr} = \overline{E}_{v.n} g(\omega)\,d\omega.$$

En intégrant cette expression sur tout le spectre de vibrations normales, c'est-à-dire dans les limites de 0 à ω_D, on obtient l'énergie des vibrations thermiques du réseau cristallin d'un solide:

$$E_{cr} = \int_0^{\omega_D} \overline{E}_{v.n} g(\omega)\,d\omega = \int_0^{\omega_D} \frac{\hbar\omega \cdot 9N\omega^2\,d\omega}{\omega_D^3 (e^{\hbar\omega/(kT)} - 1)}. \tag{3.32}$$

Ici, on a utilisé l'expression (3.22) pour $g(\omega)$ et (3.28) pour $\overline{E}_{v.n}$. La chaleur massique s'obtient en dérivant l'expression (3.32) par rapport à la température.

Dans le domaine des hautes températures où $kT > \hbar\omega_D$, on a pour toute vibration normale $E_{v.n} = kT$ et $c_V = 3Nk$, c'est-à-dire que la loi de Dulong et Petit est de nouveau confirmée.

Aux basses températures lorsque $kT \ll \hbar\omega_D$, l'intégration de (3.32) donne

$$E_{cr} = \frac{3\pi^4}{5} NkT_D \left(\frac{T}{T_D}\right)^4 \sim T^4, \tag{3.33}$$

où $T_D = \hbar\omega_D/k$. En dérivant (3.33) par rapport à la température, on trouve

$$c_V = \frac{12\pi^4}{5} Nk \left(\frac{T}{T_D}\right)^3 \sim T^3. \tag{3.34}$$

Sur la figure 3.10 la ligne continue représente la courbe de variation théorique de la chaleur massique des solides en fonction de la température et les points indiquent les données expérimentales pour

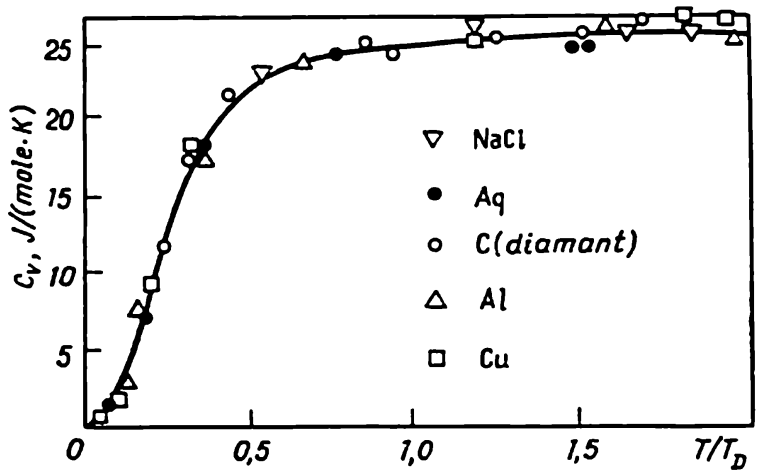

Fig. 3.10. Variation de la chaleur massique des solides en fonction de la température d'après Debye (courbe en traits pleins) et données expérimentales pour certains matériaux.

l'argent, le diamant, l'aluminium, le cuivre et le sel gemme. L'accord entre la théorie et l'expérience est tout à fait satisfaisant non seulement du point de vue qualitatif, mais aussi du point de vue quantitatif.

Variation de la concentration et de l'énergie moyenne des phonons avec la température. De (3.29) il résulte que dans le domaine des hautes températures le nombre moyen de phonons dans chaque mode croît proportionnellement à la température. Par conséquent, le nombre total de phonons n_{phon} dans le volume unité du cristal (c'est-à-dire leur concentration) est lui aussi proportionnel à T:

$$n_{phon} \sim T. \tag{3.35}$$

Quant à l'énergie moyenne du phonon, elle ne dépend pas de T:

$$\langle E_{phon} \rangle = E_{cr}/n_{phon} = \text{const}. \tag{3.36}$$

Ici, les chevrons représentent la moyenne prise sur la collectivité.

Aux basses températures les calculs montrent que

$$n_{phon} \sim T^3; \quad \langle E_{phon} \rangle \sim T. \tag{3.37}$$

Expliquons la cause d'une telle dépendance. La figure 3.11, *a* montre la courbe de variation en fonction de la température du quotient de l'énergie de la vibration normale selon Einstein (3.28) par la valeur classique kT de cette énergie. D'une façon grossièrement approchée (ligne en traits interrompus de la figure 3.11, *a*), cette dépendance peut être interprétée ainsi : aux températures $kT > \hbar\omega_{v.n}$ la vibration normale est excitée ; aux températures

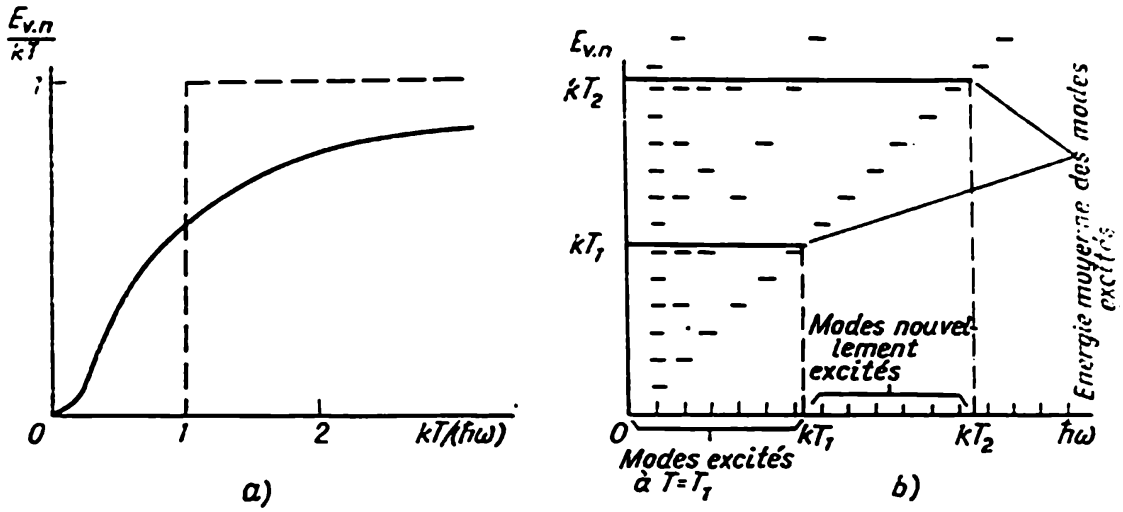

Fig. 3.11. Excitation des vibrations normales du réseau dans le domaine des basses températures :

a — variation en fonction de la température du rapport de l'énergie de la vibration normale $E_{v.n}$ à la valeur classique kT montrant que la vibration normale de fréquence $\omega_{v.n}$ est excitée principalement à $kT > \hbar\omega_{v.n}$; à des températures plus basses elle n'est excitée que faiblement ; *b* — lorsque la température augmente de T_1 à T_2, à des modes excités à T_1 s'ajoutent des modes de fréquences plus élevées nouvellement excités.

plus basses elle est excitée faiblement. La figure 3.11, *b* représente les spectres énergétiques des vibrations de fréquence la plus basse ω_1 et de fréquences plus élevées égales à $2\omega_1$, $3\omega_1$, etc. (comme pour la chaîne linéaire d'atomes). L'énergie moyenne de ces vibrations est égale à kT_1 à une certaine température peu élevée T_1, à kT_2 à la température T_2 ; à T_1 sont excités des modes de $\hbar\omega < kT_1$ et à T_2 des modes de $\hbar\omega < kT_2$. On voit sur cette figure que dans les modes qui sont déjà excités pour $T = T_1$ le nombre de phonons est doublé par suite de l'augmentation de la température et en outre il se produit l'excitation de nouveaux types de vibrations de fréquences plus élevées et d'énergies des phonons plus grandes (entre kT_1 et kT_2). C'est pourquoi $\langle E_{phon} \rangle$ croît avec la température tandis que n_{phon} croît plus vite que T. Pour $T > T_D$ tous les types possibles de vibrations sont déjà excités et une croissance ultérieure de T n'a pour

effet qu'une augmentation proportionnelle du nombre de phonons dans chaque mode.

Soulignons une fois de plus que l'écart de la dépendance $c_V(T)$ à la loi classique est lié au caractère quantique de l'excitation des vibrations, à ce que la vibration de fréquence ω ne peut pas apparaître progressivement mais exige pour son excitation d'emblée une portion d'énergie $\hbar\omega$. Aussi, aux basses températures, les vibrations de haute fréquence qui exigent de grandes portions d'énergie d'excitation sont-elles pratiquement absentes. Regardons par exemple comment arrive à un état d'équilibre thermique un cristal ventillé par un gaz inerte froid. Sur la surface du cristal il se produit un échange d'énergie entre les atomes vibrants du cristal et les atomes de gaz tombant sur le cristal. Dans ces conditions, les types de basses fréquences des vibrations peuvent être excités pratiquement par tout atome de gaz tandis que les vibrations de hautes fréquences ne peuvent le faire que par des atomes possédant une énergie élevée. Mais l'énergie des modes excités à haute fréquence peut être empruntée et leurs vibrations peuvent être amorties par tout atome frappant le cristal. Un tel échange « inégal » d'énergie entre les modes de haute fréquence et le gaz conduit à une faible excitation de ces modes. Pour ce qui est des modes de basse fréquence, l'échange d'énergie est pour eux « équivalent » et ils possèdent, de même que les particules de gaz, une énergie cinétique de $\frac{1}{2}kT$ par chaque degré de liberté.

§ 3.6. Dilatation thermique des solides

En étudiant les vibrations thermiques du réseau cristallin, nous avons introduit l'approximation harmonique (3.1) suivant laquelle la force qui prend naissance lorsque la particule se déplace de sa position d'équilibre, est proportionnelle au déplacement. La théorie de la chaleur massique de réseau des corps solides construite dans cette hypothèse s'accorde de manière suffisamment satisfaisante avec l'expérience. Mais il s'est avéré impossible d'expliquer dans le cadre de l'approximation harmonique certains phénomènes bien connus tels que par exemple la dilatation thermique des solides et leur conductibilité thermique.

En effet, reportons-nous à la courbe de variation de l'énergie potentielle U des particules du solide en fonction de la distance r qui les séparent l'une de l'autre (fig. 3.12). Si les conditions (3.1) sont satisfaites, cette courbe est une parabole $a'bc'$ dont l'équation est de la forme

$$U(r - r_0) = U(u) = -U_0 + {}^1/_2\, \beta u^2. \qquad (3.38)$$

Ici on a noté $u = r - r_0$. Pour $T = 0$ K, les particules se disposent à des distances r_0 correspondant au minimum d'énergie d'interac-

tion U (c'est-à-dire au fond du puits de potentiel $a'bc'$). Ces distances déterminent la dimension du corps à $T = 0$ K. Lorsque la température augmente, les particules commencent à vibrer autour de leurs positions d'équilibre r_0. Pour simplifier, supposons que la particule *1* est immobile et que seule la particule *2* vibre. La particule en vibration possède une énergie cinétique qui atteint sa valeur maximale E_{osc} à l'instant où elle passe par la position d'équilibre r_0. Sur la figure 3.12, l'énergie E_{osc} est portée vers le haut à partir du fond du puits de potentiel. Lorsque la particule *2* se déplace vers la gauche à partir de la position d'équilibre, l'énergie cinétique est dépensée pour vaincre les forces de répulsion par la particule *1* et se transforme en énergie potentielle d'interaction de la particule. Le

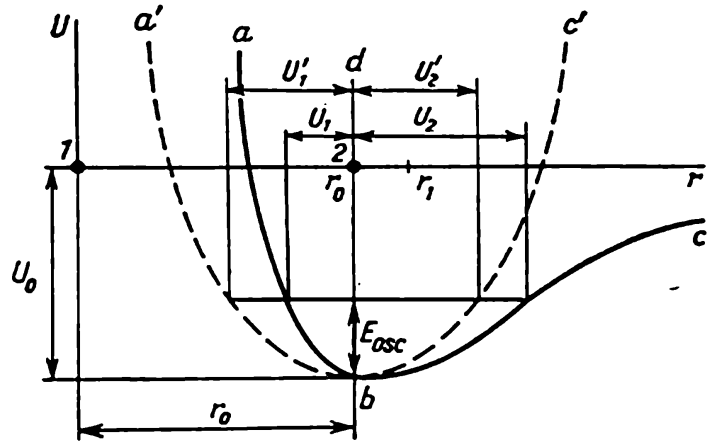

Fig. 3.12. Variation de l'énergie potentielle d'interaction des particules du corps en fonction de la distance qui les sépare (courbe en traits pleins).

mouvement vers la gauche a lieu tant que toute l'énergie cinétique de la particule E_{osc} ne se transforme en énergie potentielle. Le mouvement de la particule *2* vers la droite de la position d'équilibre s'effectue de la même manière. Comme la parabole (3.38) est symétrique par rapport à la verticale bd passant par le point r_0, les écarts maximaux de la particule *2* vers la droite (u_2') et vers la gauche (u_1') sont égaux.

Or, en réalité la courbe potentielle abc est, comme il est visible sur la figure 3.12, asymétrique par rapport au point r_0 : sa branche de gauche ba est plus fortement ascendante que sa branche de droite bc. Cela signifie que les vibrations des particules dans le solide sont anharmoniques. Pour prendre en compte l'asymétrie de la courbe potentielle on peut introduire dans l'équation (3.38) un terme supplémentaire $-\gamma u^3/3$ (γ est un coefficient de proportionnalité). Alors, les expressions (3.38) et (3.1) prennent la forme suivante :

$$U(r - r_0) = U(u) = -U_0 + {}^1/_2\, \beta u^2 - {}^1/_3\, \gamma u^3, \qquad (3.39)$$

$$F(r - r_0) = F(u) = -\frac{\partial U}{\partial u} = -\beta u + \gamma u^2. \qquad (3.40)$$

Dans le cas où la particule *2* s'écarte vers la droite ($u > 0$), le terme $\gamma u^3/3$ se retranche de $\beta u^2/2$ et la branche *bc* devient moins abrupte que la branche *bc'* ; si la particule *2* s'écarte vers la gauche ($u < 0$), le terme $\gamma u^3/3$ s'ajoute à $\beta u^2/2$ et la branche *ba* devient plus abrupte que la branche *ba'*.

Le caractère asymétrique de la courbe potentielle a pour effet que les écarts de la particule *2* vers la droite (u_2) et vers la gauche (u_1) ne sont pas égaux : cette particule s'écarte plus fortement vers la droite que vers la gauche (fig. 3.12). Ceci étant, la position moyenne de la particule *2* (point r_1) ne coïncide plus avec la position d'équilibre r_0 mais se trouve déplacée vers la droite. Cela correspond à une augmentation de $\bar{u}$ de la distance moyenne entre les particules.

Ainsi, lorsque le corps est échauffé, les distances moyennes entre les particules doivent augmenter et le corps doit se dilater. La cause en est le caractère anharmonique des vibrations des particules du corps solide dû à l'asymétrie de la courbe de variation de l'énergie d'interaction des particules en fonction de la distance qui les sépare.

Evaluons le coefficient de dilatation thermique α. La valeur moyenne de la force qui prend naissance lorsque la particule *2* s'écarte de la position d'équilibre est égale à

$$\overline{F} = -\beta\bar{u} + \gamma\overline{u^2}. \tag{3.41}$$

Lorsque la particule effectue des vibrations libres, $\overline{F} = 0$ et donc $\gamma\overline{u^2} = \beta\bar{u}$. On en tire

$$\bar{u} = \gamma\overline{u^2}/\beta. \tag{3.42}$$

L'énergie potentielle de la particule est déterminée par la relation (3.38) à une quantité du deuxième ordre de petitesse près. Lorsque les vibrations commencent, cette énergie augmente en moyenne de

$$\overline{\Delta U} = \overline{U}_{osc} = {}^1/_2\beta\overline{u^2}. \tag{3.43}$$

Comme la valeur moyenne de l'énergie potentielle de l'oscillateur est égale à la valeur moyenne de son énergie cinétique ou à la moitié de son énergie totale E_{osc}, on a

$$\beta\overline{u^2} = E_{osc}; \tag{3.44}$$

$$\bar{u} = \frac{\gamma\overline{u^2}}{\beta} = \frac{\gamma}{\beta^2} E_{osc} \tag{3.45}$$

La dilatation linéaire relative qui représente le rapport de la variation de la distance moyenne $\bar{u}$ entre les particules à la distance normale r_0 entre elles a pour valeur

$$\frac{\bar{u}}{r_0} = \frac{\gamma}{\beta^2 r_0} E_{osc},$$

et le coefficient de température de dilatation thermique est égal à

$$\alpha = \frac{1}{r_0}\frac{d\bar{u}}{dT} = \frac{\gamma}{\beta^2 r_0}\frac{dE_{osc}}{dT} = \frac{\gamma}{\beta^2 r_0}\frac{c_V}{3N}, \tag{3.46}$$

où $c_V/(3N)$ est la chaleur massique rapportée à un degré de liberté de la particule. Ainsi il se trouve que le coefficient de température de dilatation thermique est proportionnel à la chaleur massique du

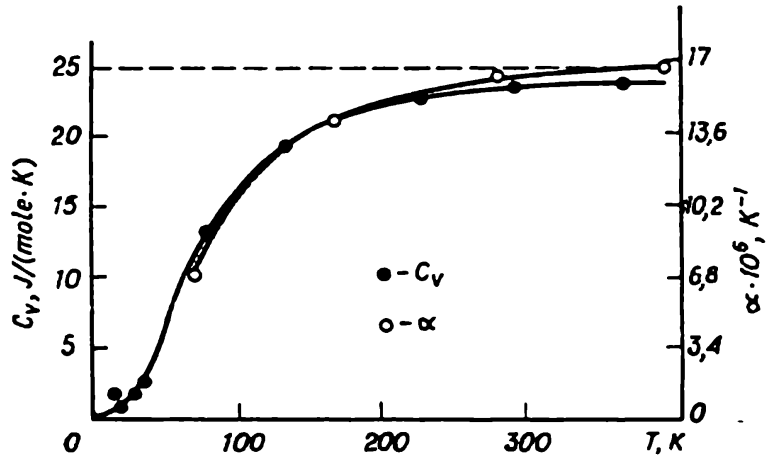

Fig. 3.13. Variation du coefficient de température de dilatation thermique α et de la chaleur massique c_V du cuivre en fonction de la température.

corps. La figure 3.13 montre à titre d'exemple la variation du coefficient de température de dilatation thermique et de la chaleur massique du cuivre en fonction de la température qui confirme le lien qui existe entre α et c_V.

Dans le domaine des hautes températures, l'énergie des vibrations de la particule par un degré de liberté, c'est-à-dire par la vibration le long d'un des axes, est égale en moyenne à kT et la chaleur massique c_V rapportée à une particule est égale à k. Ceci étant, le coefficient de dilatation thermique d'une chaîne linéaire d'atomes a pour expression

$$\alpha = \frac{\gamma k}{\beta^2 r_0}. \tag{3.47}$$

L'introduction des valeurs numériques de γ, k, β et r_0 pour les divers corps solides donne pour α des valeurs de l'ordre de 10^{-4} à 10^{-5}, ce qui s'accorde de manière satisfaisante avec l'expérience. L'expérience confirme aussi que dans le domaine des hautes températures le coefficient α ne dépend pratiquement pas de la température (fig. 3.13).

Dans le domaine des basses températures, le comportement de α est analogue à celui de c_V : il diminue lorsque la température baisse et tend vers zéro lorsque $T \to 0$ K.

Notons avant de clore ce paragraphe qu'une formule analogue à (3.46) a été proposée pour la première fois pour les métaux par Gruneisen sous la forme

$$\alpha = \frac{\delta \varkappa}{3V} c_V, \tag{3.48}$$

où $\varkappa$ est le coefficient de compressibilité du métal ; V le volume occupé par un atome ; δ la constante de Gruneisen qui prend pour divers métaux les valeurs de 1,5 à 2,5.

§ 3.7. Conductibilité thermique de réseau des solides

Un deuxième effet dû au caractère anharmonique des vibrations des atomes est la résistance thermique des solides. Elle ne pourrait pas apparaître si les atomes effectuaient des vibrations rigoureusement harmoniques, se propageant dans le réseau sous la forme d'un système d'ondes élastiques non interagissantes entre elles. L'absence d'interaction entre les ondes leur permettrait de se propager dans le cristal sans subir de diffusion, c'est-à-dire sans rencontrer aucune résistance à la manière de la lumière se propageant dans le vide.

Si l'on crée dans un tel cristal une différence de température, les atomes situés à l'extrémité chaude, dont les amplitudes de la vibration sont grandes, doivent transmettre leur énergie aux atomes voisins. Dans ce cas idéalisé, le front d'onde thermique se propagerait le long du cristal à la vitesse du son. Comme cette onde ne rencontrerait aucune résistance, la conductivité thermique d'un tel cristal serait très grande.

Dans les cristaux réels, à des températures pas trop basses, les vibrations des atomes ont un caractère anharmonique dont tient compte le deuxième terme de (3.40). L'apparition de l'anharmonicité conduit à ce que les ondes élastiques perdent leur caractère indépendant et, qu'en se rencontrant, elles entrent en interaction, échangent de l'énergie et changent la direction de leur propagation (en subissant la diffusion l'une sur l'autre). C'est précisément grâce à ces processus d'interaction des ondes élastiques que deviennent possibles le transfert d'énergie des vibrations d'une fréquence aux vibrations d'une autre fréquence et l'établissement d'un équilibre thermique dans le cristal.

Expliquons pourquoi la diffusion des ondes élastiques, l'une sur l'autre, a son origine dans l'anharmonicité des vibrations. Le fait est que la non-linéarité de la variation de la force en fonction du déplacement (3.40) signifie que le module d'élasticité dépend du déplacement des atomes depuis la position d'équilibre. Aussi, une région du cristal occupée par une onde est-elle considérée par une autre onde comme une région ayant d'autres caractéristiques élastiques, c'est-à-dire comme une hétérogénéité.

Pour décrire le processus de diffusion des vibrations normales l'une sur l'autre, il est commode d'utiliser le langage de phonons en considérant un cristal thermiquement excité comme une boîte remplie de phonons. A l'approximation harmonique, dans laquelle les vibrations normales du réseau sont indépendantes, les phonons constituent un gaz parfait (*gaz de phonons non interagissants*). Le passage aux vibrations anharmoniques est équivalent à l'introduction de l'interaction entre les phonons, qui peut avoir pour résultat la désintégration d'un phonon en deux ou plusieurs phonons et la formation d'un phonon à partir de deux. Il convient d'appeler ces processus *diffusion phonon-phonon*.

En considérant les processus de diffusion des phonons comme des collisions des particules, il est commode de se représenter le phonon comme un paquet d'ondes se déplaçant dans le cristal. En réalité, le phonon, en tant que quantum d'énergie de la vibration normale, est étendu à tout le cristal. Quant au paquet d'ondes, il est constitué par une superposition de vibrations normales. En particulier il ne peut pas être caractérisé par un seul vecteur d'onde **k**. Toutefois, si le paquet n'est pas trop étroit, la gamme de vecteurs d'onde qui le constituent n'est pas grande. Dans de tel cas, on peut parler du vecteur d'onde **k** du paquet d'ondes phononique et de l'énergie $\hbar\omega$ (**k**) de ce paquet. Pour ce qui est de la propagation de la chaleur dans l'espace, on peut la représenter comme un mouvement par diffusion des paquets d'ondes phononiques qui échangent, l'un avec l'autre, de l'impulsion et de l'énergie.

Dans la théorie cinétique des gaz on montre que le coefficient de conductibilité thermique du gaz a pour expression

$$\lambda = {}^1/_3\, lvc_V, \tag{3.49}$$

où l est la longueur de libre parcours moyen des molécules; v la vitesse de l'agitation thermique des molécules; c_V la chaleur massique par unité de volume du gaz.

Relativement au gaz phononique, il convient d'entendre par c_V la chaleur massique du volume unité du cristal, par l la longueur du libre parcours moyen du phonon, par v la vitesse (de groupe) des paquets d'ondes phononiques. D'une manière approximative, on peut considérer que la vitesse des phonons ne dépend pas de leurs vecteurs d'onde et est égale à la vitesse de propagation des vibrations de basse fréquence du réseau cristallin, c'est-à-dire à la vitesse du son.

Analysons la variation du coefficient de conductibilité thermique en fonction de la température déterminée par la formule (3.49).

Dans le domaine des hautes températures c_V ne dépend pratiquement pas de T. Quant à la longueur du libre parcours moyen des phonons, elle doit être inversement proportionnelle à la concentration des phonons si ceux-ci se diffusent l'un sur l'autre. Comme $n_{phon} \sim T$, le coefficient de conductibilité thermique doit être inversement

proportionnel à la température thermodynamique, ce qui s'accorde qualitativement avec l'expérience.

A des températures inférieures à la température de Debye, la concentration de phonons diminue brusquement lorsque T baisse de sorte que la longueur de leur libre parcours moyen augmente brusquement et atteint une valeur comparable à la dimension du cristal. Vu que généralement les parois du cristal réfléchissent les phonons dans une direction quelconque (réflexion diffuse), une baisse ultérieure de la température ne fait pas croître l parce que cette dernière se détermine tout simplement par les dimensions du cristal. La variation avec la température de la conductibilité thermique du réseau dans cette gamme de températures est déterminée par la variation de la chaleur massique c_V du cristal en fonction de T. Puisque dans le domaine des basses températures $c_V \sim T^3$, la conductibilité thermique λ doit être elle aussi proportionnelle à T^3:

$$\lambda \sim T^3. \qquad (3.50)$$

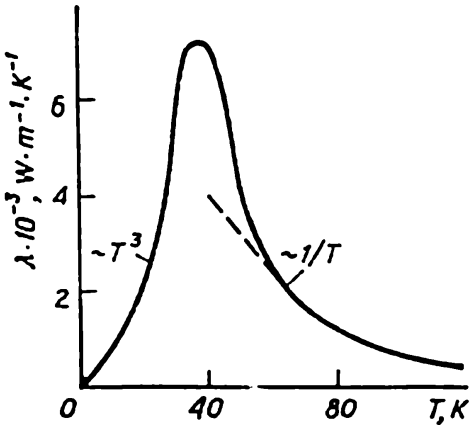

Fig. 3.14. Variation de la conductibilité thermique du saphir en fonction de la température.

Cette conclusion est, elle aussi, confirmée qualitativement par l'expérience. La figure 3.14 montre à titre d'exemple la variation de la conductibilité thermique du saphir artificiel en fonction de la température. Dans le domaine des basses températures c_V est réellement approximativement proportionnelle à T^3.

Une situation analogue doit se présenter dans des diélectriques amorphes dont les dimensions des régions de structure régulière sont, d'après leur ordre de grandeur, comparables aux dimensions atomiques. La diffusion des phonons sur les frontières de telles régions doit prédominer à toutes les T et de ce fait l ne doit pas dépendre de T. Ceci étant, le coefficient de conductibilité de tels diélectriques doit être proportionnel à T^3 dans le domaine des basses températures et être indépendant de T dans le domaine des hautes températures, ce qui est observé expérimentalement.

Théorie quantique de la conductibilité thermique. La théorie montre que lors de la diffusion des phonons doivent être observées à la fois la loi de la conservation de l'énergie et la loi de la conservation de l'impulsion des phonons. Le rôle de l'impulsion du phonon est joué par la grandeur $\hbar\mathbf{k}$. Par exemple, si deux phonons d'énergies $\hbar\omega_1$ et $\hbar\omega_2$ et d'impulsions $\hbar\mathbf{k}_1$ et $\hbar\mathbf{k}_2$ forment par suite de la collision un phonon, l'énergie et l'impulsion de ce

dernier doivent être égales à $\hbar\omega_3 = \hbar\omega_1 + \hbar\omega_2$ et $\hbar\mathbf{k}_3 = \hbar\mathbf{k}_1 + \hbar\mathbf{k}_2$ respectivement.

L'addition de deux phonons de basses fréquences se propageant dans une même direction peut être observée expérimentalement : en excitant deux ondes de fréquences ω_1 et ω_2 à une extrémité d'une barre cristalline, on peut observer à l'autre extrémité une vibration de fréquence $\omega_3 = \omega_1 + \omega_2$.

Dans le cas général, l'exigence concernant l'observation simultanée des lois de la conservation de l'énergie et de l'impulsion lors de l'interaction des phonons est assez sévère. Par exemple, l'addition

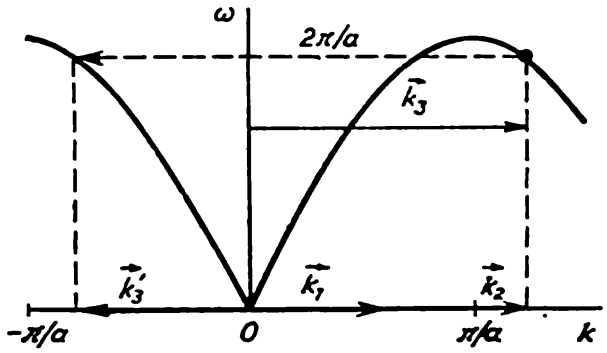

Fig. 3.15. Explication de l'influence des processus de « transfert » sur le mouvement des phonons.

de deux phonons de grande énergie est impossible si le nouveau phonon doit avoir une énergie supérieure à la valeur maximale possible $\hbar\omega_{max}$. En même temps, c'est précisément l'interaction des phonons de grandes énergies qui doit conduire à une limitation de la conductibilité thermique. Expliquons ceci sur l'exemple de la chaîne linéaire d'atomes.

L'addition de deux phonons de petites impulsions donne un nouveau phonon qui se propage dans la même direction que les deux phonons disparus et possède une énergie égale à la somme des énergies des phonons disparus. Ainsi, le flux d'énergie est conservé. Mais l'addition de deux phonons d'impulsions $\hbar k_{1,2} > \pi/(2a)$ donne un phonon d'impulsion $\hbar k_3 > \pi/a$ qui se situe en dehors de la première zone de Brillouin. C'est ce qu'on appelle « processus de transfert » ou processus U (de l'allemand *Umklapprozesse*).

Comme il a été montré au § 3.1, toute vibration dont le vecteur d'onde se situe en dehors de la première zone de Brillouin peut être décrite avec le même succès à l'aide d'une onde dont le vecteur d'onde diffère de la quantité $2\pi/a$ ou de son multiple et tombe dans la première zone de Brillouin (flèche en traits interrompus de la figure 3.15). Autrement dit, l'impulsion $\hbar\mathbf{k}$ du phonon est définie à tout vecteur du réseau réciproque près (v. § 1.3). C'est pourquoi la quantité $\hbar\mathbf{k}$ est souvent appelée *quasi-impulsion du phonon* et non pas impulsion.

Comme il est visible sur la figure 3.15, la vitesse de groupe du paquet d'ondes d'impulsion $\hbar k_3' = \hbar k_3 - 2\pi/a$ est dirigée dans le sens opposé à $\hbar k_1$ et $\hbar k_2$. (La vitesse de groupe est définie comme étant égale à $d\omega/dk$ pour des valeurs considérées de k). Ainsi, une telle interaction devrait conduire au changement de la direction du flux d'énergie. Pourtant, comme il a été dit plus haut, lors de l'addition de tels phonons il est impossible de respecter la loi de la conservation de l'énergie.

Dans un cristal tridimensionnel où il existe plusieurs polarisations d'ondes et l'énergie des phonons dépend non seulement du module du vecteur d'onde mais aussi de son orientation, les possibilités de diffusion des phonons sont plus grandes. En outre, il convient de prendre en considération des processus faisant intervenir un grand nombre de phonons à la fois.

Ceci étant, la théorie quantique de la conductibilité thermique s'est avérée très complexe et, pour le moment, elle n'est pas en mesure de prédire non seulement une valeur exacte mais même l'ordre de grandeur de la conductibilité thermique λ du réseau cristallin.

CHAPITRE 4

ÉLÉMENTS DE THÉORIE DES BANDES DES SOLIDES

La théorie électronique moderne des solides est fondée sur la théorie des bandes. C'est seulement sur la base de cette théorie qu'il est devenu possible d'expliquer les propriétés électriques des métaux, des diélectriques et des semiconducteurs.

§ 4.1. Niveaux d'énergie électroniques dans un atome isolé

Au début du XXe siècle les physiciens se sont rendus compte que les phénomènes atomiques ne pouvaient être décrits ni comme un mouvement des particules, ni comme des processus purement ondulatoires. Les particules microscopiques — les électrons, les protons, les atomes, etc., réunissent organiquement les propriétés corpusculaires et les propriétés ondulatoires. Elles sont toujours enregistrées comme des particules, mais leur mouvement est décrit par l'*équation d'onde de Schrödinger*. Pour une microparticule de masse m en mouvement dans un champ de force et possédant dans ce champ une énergie potentielle $U(x, y, z, t)$ l'équation de Schrödinger est de la forme

$$-i\hbar \frac{\partial \Psi}{\partial t} = \frac{\hbar^2}{2m}\left(\frac{\partial^2 \Psi}{\partial x^2} + \frac{\partial^2 \Psi}{\partial y^2} + \frac{\partial^2 \Psi}{\partial z^2}\right) - U(x, y, z, t)\Psi. \quad (4.1)$$

La fonction $\Psi(x, y, z, t)$, solution de cette équation, est appelée *fonction d'onde*. L'interprétation physique de cette fonction est la suivante : le produit de Ψ par la fonction Ψ^*, qui est sa conjuguée complexe, et par le volume dV est la probabilité pour que la particule soit localisée à l'instant t dans l'élément de volume donné dV.

L'énergie potentielle U figurant dans l'équation de Schrödinger est dans le cas général fonction des coordonnées et du temps. Toutefois, pour de nombreux problèmes pratiques importants, elle n'est que fonction des coordonnées et ne dépend pas du temps. Dans ce cas il existe une collection de solutions indépendantes de l'équation de Schrödinger pour lesquelles $\Psi\Psi^*$ est indépendant du temps. De tels états de microparticules sont dits *stationnaires*.

La détermination des états stationnaires est un problème très important parce que tout autre état, y compris l'état non stationnaire, peut être représenté comme une superposition de fonctions d'onde stationnaires. Malheureusement, une solution rigoureuse de

l'équation de Schrödinger n'est obtenue que pour certains systèmes simples et notamment pour l'atome d'hydrogène. La résolution de l'équation de Schrödinger pour un électron dans un champ coulombien du proton donne une multitude d'états stationnaires possibles. Leur énergie, comptée à partir de l'énergie d'un électron libre, non lié au noyau, est donnée par la formule

$$E_{nl} = -\frac{e^4 m_e}{8\varepsilon_0^2 h^2}\,\frac{1}{n^2}\,, \tag{4.2}$$

où ε_0 est la permittivité du vide ; h la constante de Planck ; e et m_e sont la charge et la masse de l'électron ; $n = 1, 2, 3, \ldots$ est un nombre dit *quantique principal.*

L'électron unique dans l'atome d'hydrogène occupe l'état fondamental le plus bas ($n = 1$). On peut le faire passer à des niveaux énergétiques plus élevés, en lui transmettant de l'extérieur une portion correspondante d'énergie. Dans les atomes à plusieurs électrons chaque électron interagit non seulement avec le noyau mais aussi avec d'autres électrons. Il en résulte une forte complication de l'équation de Schrödinger, ce qui ne permet pas d'obtenir les solutions exactes. Ainsi, on a recours à des méthodes de résolution approchée.

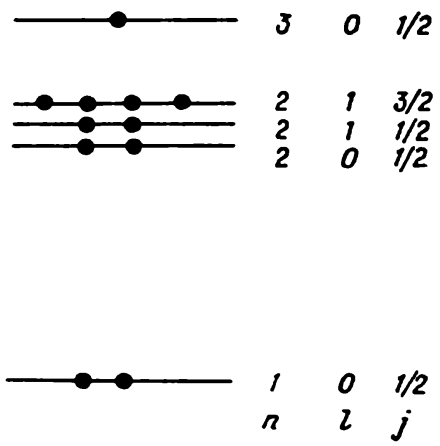

Fig. 4.1. Niveaux d'énergie des électrons dans un atome à plusieurs électrons.

La fonction d'onde du système d'électrons qui dépend maintenant des coordonnées de tous les électrons de l'atome est représentée sous la forme d'une combinaison de fonctions dont chacune ne dépend des coordonnées que d'un seul électron. La forme de chacune de ces fonctions et l'énergie de l'électron dans l'atome qui lui correspond se déterminent non seulement par le nombre quantique principal n, comme dans l'atome d'hydrogène, mais aussi par le moment orbital $\mathbf{l}$ et par le moment total $\mathbf{j}$ qui est le résultat de l'addition vectorielle du moment orbital et du moment de spin. La figure 4.1 montre (non à l'échelle) le schéma des niveaux d'énergie d'un atome comportant onze électrons (Na).

Une même énergie peuvent avoir dans l'atome deux (ou plusieurs) électrons de différentes fonctions d'onde. En d'autres termes, plusieurs niveaux énergétiques peuvent coïncider. De tels niveaux sont dits *dégénérés.* C'est ainsi par exemple que le niveau de $j = \hbar/2$ est doublement dégénéré. Un tel niveau peut être occupé par deux électrons dont les fonctions d'onde diffèrent par la projection du moment

total j sur une direction donnée. Ces projections (en unités de $\hbar$) peuvent prendre des valeurs de $-j$ à $+j$ avec des intervalles de 1. Ainsi, pour $j = 1/2$ ces projections peuvent être égales à $-1/2$ ou à $+1/2$. Pour $j = 3/2$, elles peuvent prendre les valeurs de $-3/2$, $-1/2$, $1/2$, $3/2$. Cela signifie que le niveau de $j = 1/2$ peut contenir deux électrons, ce niveau est doublement dégénéré; le niveau avec $j = 3/2$ est dégénéré d'ordre 4, et ainsi de suite (fig. 4.1).

§ 4.2. Mise en commun des électrons dans le cristal

Dans les solides, les distances entre les atomes sont si petites que tout au moins les électrons de valence de chaque atome se trouvent soumis à un champ suffisamment fort des atomes voisins. Si l'on considère en première approximation que les noyaux des atomes du solide sont immobiles dans les nœuds du réseau cristallin, l'équation de Schrödinger pour la détermination des états stationnaires du système d'électrons de valence prendra la forme suivante:

$$-\frac{\hbar}{2m}\sum_i\left(\frac{\partial^2\Psi}{\partial x_i^2}+\frac{\partial^2\Psi}{\partial y_i^2}+\frac{\partial^2\Psi}{\partial z_i^2}\right)+$$
$$+U(x_1,\ y_1,\ z_1,\ x_2,\ y_2,\ z_2,\ \ldots)\,\Psi=E\Psi, \qquad (4.3)$$

où x_i, y_i, z_i sont les coordonnées du i-ième électron et E est l'énergie de tout le système électronique. La fonction d'onde du système Ψ dépend des coordonnées de tous les électrons:

$$\Psi = \Psi\,(x_1,\ y_1,\ z_1,\ x_2,\ y_2,\ z_2,\ \ldots). \qquad (4.4)$$

L'expression de l'énergie potentielle U comprend non seulement l'énergie d'interaction de chaque électron dans le champ de tous les ions du réseau, mais aussi l'énergie d'interaction coulombienne des électrons l'un avec l'autre. Une équation si compliquée comme (4.3) ne peut pas être résolue sous la forme générale. De même que dans le cas d'un atome à plusieurs électrons, on s'efforce d'obtenir la fonction d'onde du système électronique en combinant les fonctions d'onde ψ_i dépendant uniquement des coordonnées d'un seul électron et constituant les solutions des équations pour un seul électron:

$$-\frac{\hbar^2}{2m}\left(\frac{\partial^2\psi_i}{\partial x^2}+\frac{\partial^2\psi_i}{\partial y^2}+\frac{\partial^2\psi_i}{\partial z^2}\right)+U_i\,(x,\ y,\ z)\,\psi_i=E_i\psi_i, \qquad (4.5)$$

où E_i est l'énergie du i-ième électron.

Mais même dans ce cas, que l'on appelle *approximation monoélectronique*, le problème reste très difficile à résoudre parce que l'énergie potentielle $U_i\,(x,\ y,\ z)$ de chaque électron dépend des états de tous les autres électrons. Heureusement, dans bien des cas, les con-

clusions sur le comportement des électrons dans le cristal peuvent être faites à partir des raisonnements généraux.

Nous allons considérer les électrons dans un *cristal parfait*, c'est-à-dire dans un cristal possédant une structure régulière idéale, sans défauts, ne contenant pas d'atomes étrangers et n'effectuant pas de vibrations thermiques. Dans un tel cristal, les ions produisent une répartition parfaitement périodique de potentiel. Examinons le

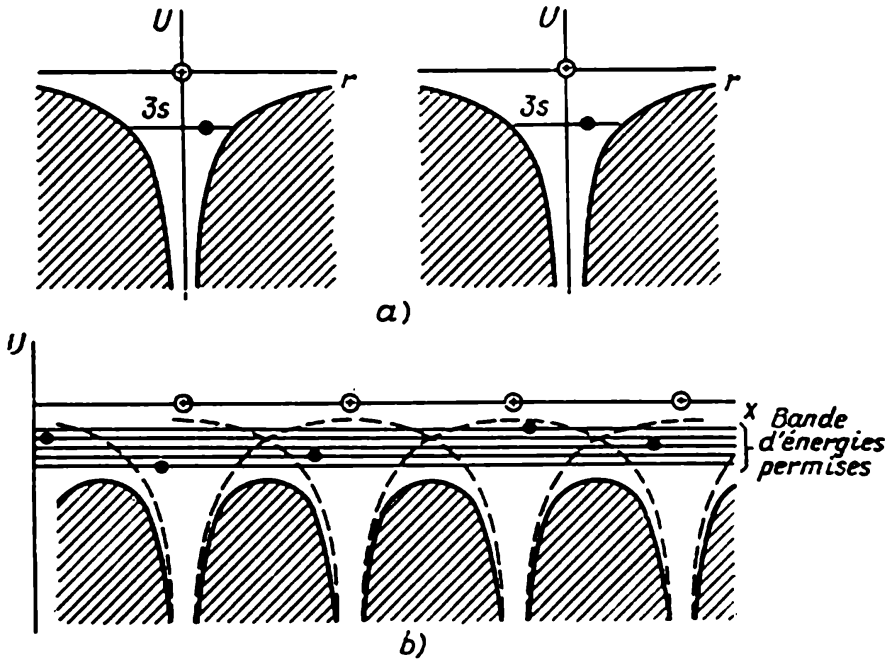

Fig. 4.2. Changement d'état des électrons lors du rapprochement des atomes: *a* — niveaux d'énergie des électrons de valence des atomes de sodium éloignés à des distances nettement supérieures au paramètre du réseau; *b* — bande d'énergie formée lors du rapprochement des atomes dans le cristal.

mouvement d'un électron de valence dans un tel champ de forces sur l'exemple de cristal de sodium.

Arrangeons N ions de sodium de telle sorte qu'ils forment un réseau spatial propre au sodium mais se trouvent à des distances si grandes l'un de l'autre que l'interaction entre eux peut être négligée. Alors, l'électron considéré peut se trouver près de n'importe lequel des ions du réseau pendant un temps infiniment long parce qu'il est maintenu par la barrière de potentiel produite par le champ coulombien de l'ion (fig. 4.2, *a*). Lorsque les ions se rapprochent jusqu'à des distances normales dans le cristal, les courbes potentielles (lignes en traits interrompus de la figure 4.2, *b*) se superposent en partie pour former des courbes résultantes montrées sur la figure 4.2, *b* en traits pleins. Dans ce cas il peut se trouver que la hauteur de la barrière entre les ions soit inférieure à la position initiale du niveau énergétique de l'électron de valence. Dans un tel cas, l'électron reçoit

la possibilité de se déplacer, pratiquement sans entraves, à l'intérieur du cristal d'un atome à un autre. Sa fonction d'onde est délocalisée sur tout le cristal. C'est ainsi que les électrons de valence sont mis en commun, chacun d'eux appartient maintenant à tout le cristal.

Si les barrières de potentiel ne descendaient pas, lors du rapprochement des atomes, au-dessous du niveau d'énergie de l'électron dans l'atome, les électrons seraient quand même mis en commun par

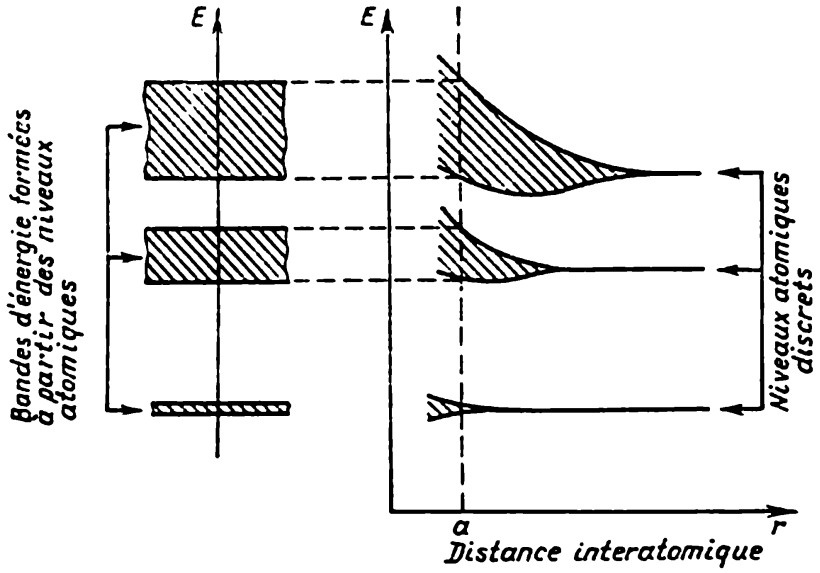

Fig. 4.3. Décomposition des niveaux d'énergie et formation de bandes d'énergie dans le cristal.

effet tunnel, c'est-à-dire la traversée de la barrière de potentiel par l'électron. Cet effet est essentiellement quantique, ce qui démontre la nature ondulatoire des microparticules.

La probabilité de traversée par effet tunnel diminue brusquement lorsque la hauteur et la largeur de la barrière augmentent. Cela signifie que les électrons occupant des niveaux plus profonds peuvent eux aussi se déplacer en principe d'un atome à un autre mais la vitesse d'un tel mouvement ne peut pas être grande.

Ainsi, les électrons mis en commun dans le cristal peuvent être assimilés dans une certaine mesure à des électrons libres. Dans le cas considéré, avant la formation du cristal, les électrons de valence 3*s* de l'atome de sodium possédaient une même énergie, mais chacun d'eux se trouvait près de « son » atome. Or, dans le solide, chacun de ces électrons est devenu appartenant à tout le cristal de sorte que de ce point de vue les électrons sont devenus indiscernables. Mais c'est de façon différente (dans des directions différentes et avec des vitesses différentes) qu'ils peuvent être maintenant animés d'un mouvement de translation dans le cristal. Ceci étant, l'énergie de divers électrons

devient différente. Un niveau énergétique étroit s'étale en *bande* d'énergies permises. Plus grandes sont les vitesses du mouvement de translation que peuvent avoir les électrons mis en commun, plus large est la bande d'énergie correspondante. C'est pourquoi, les niveaux énergétiques profonds des atomes donnent naissance à des bandes d'énergies étroites séparées par de larges intervalles d'énergies interdites. Quant aux niveaux supérieurs des atomes, ils forment de larges bandes séparées par des intervalles étroits d'énergies interdites de sorte que ces bandes peuvent même se recouvrir. Le caractère général du spectre énergétique des électrons dans un cristal est représenté sur la figure 4.3.

§ 4.3. Fonction d'onde de l'électron dans le cristal. Zones de Brillouin

Dans les limites de chaque maille élémentaire du cristal la fonction d'onde de l'électron doit être plus ou moins proche de la fonction d'onde de l'électron dans l'atome. En même temps elle doit être similaire à la fonction d'onde d'un électron libre animé d'un mouvement de translation dans l'espace. Cette dernière se présente pour l'état stationnaire sous la forme

$$\psi = Ce^{i(\mathbf{kr}-\omega t)}, \tag{4.6}$$

où C est un coefficient constant et $\mathbf{k}$ le vecteur d'onde de l'électron qui détermine sa vitesse et son énergie.

Suivant le *théorème de Bloch*, la fonction d'onde de l'état stationnaire d'un électron dans le cristal peut être représentée toujours sous la forme

$$\psi(\mathbf{r}, t) = u(\mathbf{r})\, e^{i(\mathbf{kr}-\omega t)}, \tag{4.7}$$

où $u(\mathbf{r})$ est un multiplicateur périodique de période égale à celle du réseau. Cette fonction présente une analogie évidente avec la fonction d'onde de l'électron libre (4.6) parce que les deux fonctions font intervenir le facteur $e^{i(\mathbf{kr}-\omega t)}$. La différence entre elles tient à ce que la fonction de Bloch $u(\mathbf{r})$ déterminant la probabilité pour l'électron de se trouver dans une région ou dans une autre du cristal, dépend de la coordonnée dans les limites de chaque maille élémentaire bien qu'elle se répète d'une maille à l'autre. Dans les limites de chaque maille la fonction de Bloch doit être suffisamment proche de la fonction d'onde de l'électron dans l'atome surtout pour des électrons situés dans les bandes profondes. Pour les mêmes valeurs de $\mathbf{k}$, cette fonction n'est pas la même pour les électrons situés dans les différentes bandes formées à partir des niveaux atomiques différents. Ainsi, elle doit être assortie d'indice n indiquant le numéro de la bande. D'un autre côté, les fonctions de Bloch, d'une même bande, pour des électrons ayant des valeurs différentes du vecteur

d'onde, peuvent différer l'une de l'autre. Ainsi, dans le cas général cette fonction dépend encore du paramètre **k**:

$$u(\mathbf{r}) = u_{n,\mathbf{k}}(\mathbf{r}). \tag{4.8}$$

Alors la partie de la fonction d'onde (4.7) dépendant des coordonnées doit se récrire sous la forme

$$\psi_{n,\mathbf{k}}(\mathbf{r}) = u_{n,\mathbf{k}}(\mathbf{r})\, e^{i\mathbf{kr}}. \tag{4.9}$$

Première zone de Brillouin. La division de la fonction d'onde de l'électron en deux facteurs $u(\mathbf{r})$ et $e^{i\mathbf{kr}}$ n'est pas une opération univoque. Montrons-le sur l'exemple d'une fonction d'onde unidimensionnelle de l'électron animé d'un mouvement dans la direction des x du cristal dont la période du réseau cristallin est égale à a:

$$\psi_{n,k}(x) \equiv u_{n,k}(x)\, e^{ikx}. \tag{4.10}$$

Multiplions maintenant le second membre de cette identité par $e^{-i(2\pi/a)x}\, e^{i(2\pi/a)x} \equiv 1$, il vient

$$\psi_{n,k}(x) \equiv [u_{n,k}(x)\, e^{-i(2\pi/a)x}]\, e^{i(k+2\pi/a)x}.$$

Le cofacteur entre crochets satisfait aux exigences que doivent satisfaire les fonctions de Bloch parce qu'il représente le produit de la fonction $u_{n,k}(x)$ périodique de période a par la fonction $\exp[-i(2\pi/a)x]$ également périodique de période a. Ainsi, la même fonction (4.10) est représentée maintenant comme une fonction d'onde d'un électron de vecteur d'onde $k + 2\pi/a$ au lieu de k et avec un autre multiplicateur de Bloch qu'on peut désigner maintenant comme

$$u_{n,k+2\pi/a}(x) \equiv u_{n,k}(x)\, e^{-i(2\pi/a)x},$$

et qui constitue une fonction périodique de période égale à la période a du cristal. On pourrait, en procédant exactement de la même manière, remplacer k dans (4.10) par $-2\pi/a$; $\pm 4\pi/a$; $\pm 6\pi/a$ et ainsi de suite, c'est-à-dire par un vecteur multiple du vecteur du réseau réciproque.

Cette non-univocité de la définition du vecteur d'onde de l'électron dans le cristal peut être éliminée si l'on convient de ne considérer les vecteurs d'onde que dans les limites de $-\pi/a$ à $+\pi/a$, c'est-à-dire dans les limites de la première zone de Brillouin. Une méthode analogue a été utilisée lors de l'étude des vibrations du réseau cristallin (v. §§ 3.1 et 3.3). Ainsi la collection de valeurs de **k** dans les limites de la première zone de Brillouin représente de façon exhaustive toutes les valeurs physiquement différentes des vecteurs d'onde de l'électron dans le cristal.

Courbes de dispersion. Dans chaque bande énergétique l'énergie de l'électron est fonction de son vecteur d'onde $E = E_n(\mathbf{k})$. La fi-

gure 4.4 montre les courbes de dispersion les plus simples de deux bandes d'énergie. Dans le cas général, l'énergie de l'électron dépend non seulement du module de son vecteur d'onde mais également de

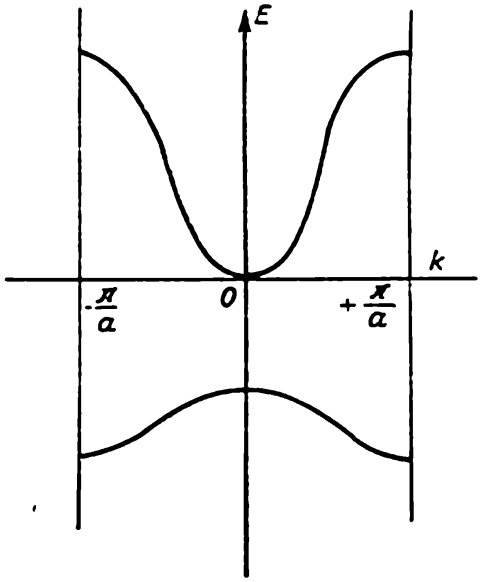

Fig. 4.4. Courbes de dispersion simples pour deux bandes d'énergie.

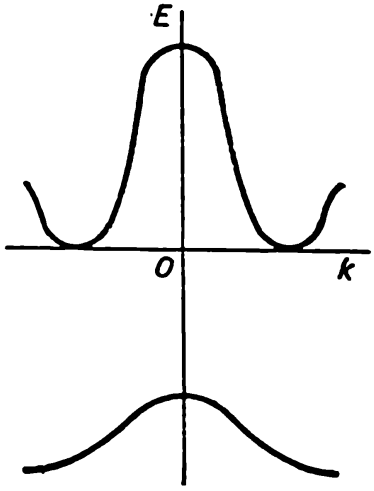

Fig. 4.5. Courbe de dispersion pour deux bandes d'énergie du silicium dans la direction [100].

l'orientation de **k** et la forme des courbes de dispersion est plus compliquée que celle des courbes de la figure 4.4. La figure 4.5 montre à titre d'exemple les courbes de dispersion pour le cristal de silicium dans la direction [100].

§ 4.4. Mouvement de translation des électrons dans le cristal

La probabilité pour que l'électron se trouve à l'instant t dans un élément de volume dV est déterminée par la quantité $\psi\psi^* dV$. Pour un électron libre à l'état stationnaire cette probabilité ne dépend, suivant (4.6), ni du temps, ni de la coordonnée:

$$\psi\psi^* dV = C^2 dV. \tag{4.11}$$

L'intégrale $\int_{(V)} \psi\psi^* dV$ étendue à tout l'espace doit être égale à l'unité parce qu'elle exprime un fait certain que la microparticule se trouve dans cet espace:

$$\int_{(V)} \psi\psi^* dV = C^2 V = 1.$$

D'où $C = 1/\sqrt{V}$.

A partir des fonctions ψ des états stationnaires (4.6) dont les vecteurs d'onde **k** ont des valeurs proches, on peut former un état non stationnaire — un paquet d'ondes se déplaçant dans l'espace de même que par exemple une impulsion de courant limitée dans le

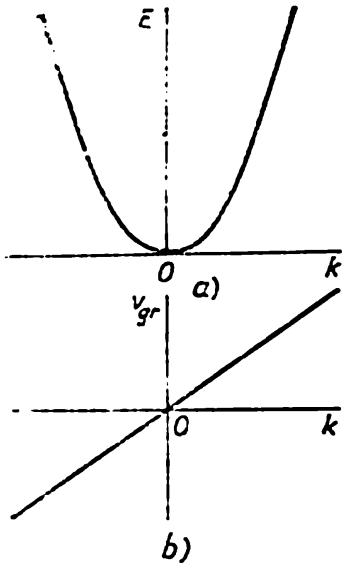

Fig. 4.6. Courbe de dispersion (*a*) et variation de la vitesse de groupe d'un électron libre en fonction du vecteur d'onde (*b*).

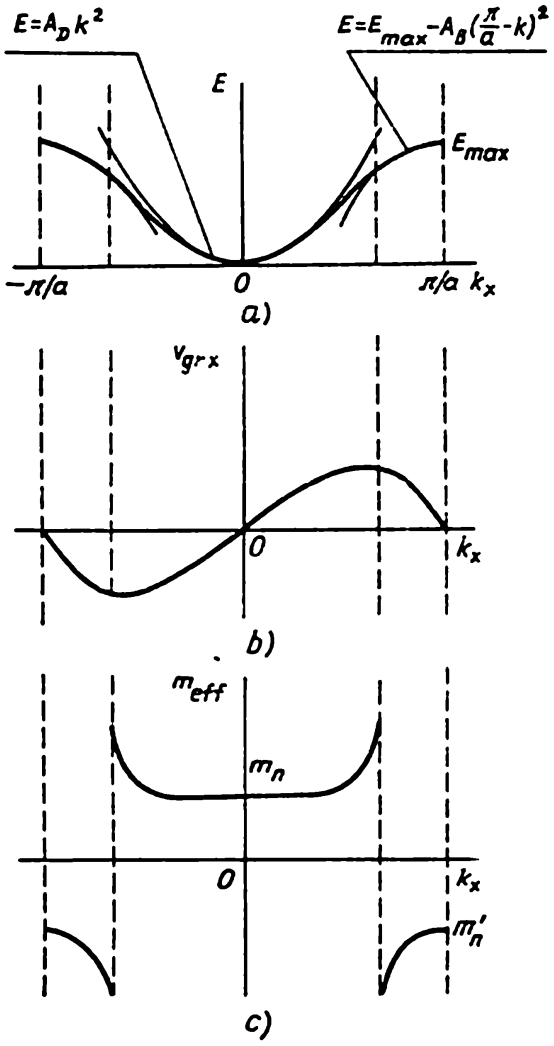

Fig. 4.7. Courbe de dispersion simple (*a*) et variation de la vitesse de groupe (*b*) et de la masse effective (*c*) d'un électron dans le cristal en fonction du vecteur d'onde.

temps peut être représentée par une superposition de sinusoïdes infinies dans le temps.

La vitesse de déplacement d'un tel paquet d'ondes dans l'espace — la *vitesse de groupe de l'électron* — est liée à son vecteur d'onde par la relation

$$\mathbf{v}_{gr} = \frac{\hbar}{m_e}\mathbf{k}, \tag{4.12}$$

où $m_e = 9{,}1 \cdot 10^{-31}$ kg est la masse de l'électron. L'énergie de l'électron libre a pour valeur

$$E = E_{cin} = \frac{m_e v_{gr}^2}{2} = \frac{\hbar^2 k^2}{2m_e}. \tag{4.13}$$

On voit que la courbe de dispersion pour l'électron libre est une parabole (fig. 4.6, *a*). La fonction $v_{gr}(k)$ pour un tel électron est représentée sur la figure 4.6, *b*.

Pour un électron dans le cristal on peut, pour lui aussi, former un paquet d'ondes à partir de la fonction d'onde (4.7) dont les vecteurs d'onde **k** ont des valeurs proches. La vitesse de son déplacement dans le cristal se détermine par l'expression

$$\mathbf{v}_{gr} = \frac{1}{\hbar} \operatorname{grad}_{\mathbf{k}} E. \tag{4.14}$$

On en tire pour la composante de la vitesse suivant l'axe des x

$$v_{gr} = \frac{1}{\hbar} \frac{\partial E}{\partial k_x}. \tag{4.15}$$

On trouve aussi des expressions analogues pour les composantes de la vitesse le long des axes y et z. (Notons qu'à partir des formules (4.14) et (4.13) on peut obtenir l'expression (4.12) pour la vitesse de groupe d'un électron libre.) La figure 4.7, *a* montre à titre d'exemple la courbe de dispersion la plus simple pour un électron dans le cristal et la figure 4.7, *b* la dépendance $v_{gr.\,x}(k_x)$.

On notera que lorsque k_x s'approche de π/a, l'énergie de l'électron augmente et sa vitesse diminue. Le fait est que l'énergie de l'électron dépend non seulement de sa vitesse de mouvement de translation dans le cristal mais également de la vitesse de mouvement de l'électron autour de l'ion dans chaque maille élémentaire. Cette énergie comprend aussi l'énergie potentielle de l'électron dans le champ électrique périodique du cristal. Toutes ces composantes de l'énergie peuvent varier en fonction de **k**, ce qui détermine justement une forme plus compliquée des courbes de dispersion pour un électron dans le cristal par rapport à l'électron dans l'espace libre.

§ 4.5. Mouvement de l'électron dans le cristal sous l'action d'une force extérieure

En l'absence de champ extérieur, les états de l'électron (4.9) dans un cristal parfait sont stationnaires quelles que soient les valeurs de son vecteur d'onde **k**, c'est-à-dire que l'électron peut séjourner dans chacun de ces états pendant un temps infiniment long.

L'application d'un champ électrique extérieur a pour effet de provoquer une variation du vecteur d'onde **k**. La loi de variation de **k** en fonction du temps peut être obtenue à partir de la loi de la conservation de l'énergie : la variation dE de l'énergie de l'électron sous

l'action d'une force extérieure F dirigée suivant la vitesse de groupe de l'électron a pour expression

$$dE = F\,dx = Fv_{gr}\,dt.$$

En faisant usage de la relation (4.14), on obtient

$$dE = \frac{F}{\hbar}\frac{dE}{dk}\,dt,$$

d'où

$$\frac{d}{dt}(\hbar k) = F \tag{4.16}$$

ou encore, sous forme vectorielle,

$$\frac{d}{dt}(\hbar \mathbf{k}) = \mathbf{F}. \tag{4.17}$$

Ces relations montrent que la valeur de $\hbar\mathbf{k}$ pour un électron dans le cristal varie avec le temps sous l'effet d'une force extérieure exactement de la même façon que l'impulsion $\mathbf{p}$ d'une particule en mécanique classique :

$$\frac{d\mathbf{p}}{dt} = \mathbf{F}.$$

Toutefois, comme il a été montré au § 4.2, à la différence d'un électron dans l'espace libre, pour un électron dans le cristal il existe une non-univocité de la détermination de la grandeur $\mathbf{k}$ elle-même (on peut la faire varier par exemple dans la direction [100] de la quantité $\pm 2\pi/a$; $\pm 4\pi/a$ et ainsi de suite, c'est-à-dire d'une quantité multiple du vecteur du réseau réciproque). Cela explique pourquoi la grandeur $\hbar\mathbf{k}$ est appelée *quasi-impulsion de l'électron.*

Déterminons maintenant l'accélération a acquise par l'électron sous l'action de la force extérieure (par raison de simplification, supposons de nouveau que la force s'exerce dans le sens de mouvement) :

$$a = \frac{dv_{gr}}{dt} = \frac{d}{dt}\left(\frac{1}{\hbar}\frac{dE}{dk}\right).$$

Comme l'énergie est une fonction composée du temps : $E = E(k(t))$, on obtient

$$a = \frac{1}{\hbar}\frac{d^2E}{dk^2}\cdot\frac{dk}{dt}.$$

En exprimant dk/dt en fonction de F, on trouve

$$a = \frac{1}{\hbar^2}\frac{d^2E}{dk^2}F. \tag{4.18}$$

Dans cette expression la grandeur $(1/\hbar)\,d^2E/dk^2$ joue le même rôle que $1/m$ (m étant la masse de la particule) dans la seconde loi de Newton. En d'autres termes, sous l'action d'une force extérieure

l'électron se déplace en moyenne dans le champ périodique du cristal comme le ferait sous l'effet de la même force un électron libre s'il possédait une masse

$$m_{eff} = \frac{\hbar^2}{d^2E/dk^2}. \tag{4.19}$$

La grandeur m_{eff} est appelée *masse effective de l'électron*. En attribuant à l'électron se trouvant dans le champ périodique du cristal la masse m_{eff}, nous pouvons considérer que cet électron est libre et décrire son mouvement dans un champ extérieur comme celui d'un électron libre ordinaire.

Cependant la masse effective, qui traduit une particularité du mouvement de l'électron dans le champ périodique du cristal, est une grandeur particulière. Tout d'abord elle peut être aussi bien positive que négative et en valeur numérique elle peut être beaucoup de fois tant plus grande que plus petite que la masse au repos m_e de l'électron. En outre, la masse effective de l'électron dans le cristal peut varier lorsque son vecteur d'onde varie. Ce qui vient d'être dit est illustré par la figure 4.7, *c*.

Pour des électrons situés près du niveau inférieur de la bande l'énergie $E_{inf} = E_{min} + A_{inf}k^2$ dont la dérivée seconde par rapport à k est égale à $d^2E/dk^2 = 2A_{inf}$. En portant cette valeur dans (4.19), on obtient l'expression suivante pour la masse effective de l'électron que nous noterons m_n :

$$m_n = \hbar^2/(2A_{inf}).$$

Comme $A_{inf} > 0$, $m_n > 0$. Ainsi, les électrons situés près du niveau inférieur de la bande d'énergie possèdent une masse effective positive. C'est pourquoi, dans le champ extérieur produit dans le cristal, ces électrons se comportent normalement en s'accélérant dans le sens de la force appliquée. La différence de tels électrons avec les électrons libres consiste en ce que leur masse effective peut différer fortement de la masse au repos. Plus large est la bande énergétique permise, plus ascendante est la courbe de dispersion, plus grand est A_{inf} et plus petite est la masse effective des électrons situés près du niveau inférieur de cette bande.

Pour des électrons situés près du sommet de la bande l'énergie $E_{som} = E_{max} - A_{som}(\pi/a - k)^2$, la dérivée seconde de E par rapport à k est $d^2E/dk^2 = -2A_{som}$ et la masse effective que nous noterons m'_n a pour valeur

$$m'_n = -\hbar^2/(2A_{som}).$$

C'est une grandeur négative. Le comportement de tels électrons dans le champ extérieur produit dans le cristal est anomal : ils sont accélérés dans un sens opposé à l'action de la force extérieure. La valeur numérique de m'_n se détermine dans ce cas par la largeur de la bande d'énergie : plus large est la bande, plus petit est $|m'_n|$.

Regardons maintenant ce qui détermine un comportement si « étrange » des électrons dans le cristal.

Pour un électron libre, tout le travail A effectué par la force extérieure est dépensé pour augmenter l'énergie cinétique du mouvement de translation

$$E_{cin} = \frac{m_e v^2}{2} = \frac{\hbar^2}{2m_e} k^2.$$

En dérivant deux fois l'énergie E_{cin} par rapport à k, on obtient $d^2E/dk^2 = \hbar^2/m_e$. En portant cette valeur dans la formule (4.19), on trouve $m_{eff} = m_e$. Ainsi, la masse effective d'un électron libre est égale tout simplement à sa masse au repos.

La situation peut être différente dans le cas d'un électron dans le cristal où il possède, en plus de l'énergie cinétique du mouvement de translation dans le cristal, encore une énergie potentielle. Lors du mouvement de l'électron sous l'action de la force F extérieure, une partie du travail effectué par cette force peut être transformée en énergie cinétique E_{cin} et une autre en énergie potentielle U si bien que $A = \Delta E_{cin} + \Delta U$. Dans ce cas l'énergie cinétique et donc la vitesse de mouvement de l'électron croissent plus lentement que dans le cas de l'électron libre. Tout se passe comme si l'électron était devenu *plus lourd* et se déplaçait sous l'effet de la force F avec une accélération inférieure à celle de l'électron libre.

Pourtant il est possible que dans le cristal se présente un cas où lors du mouvement de l'électron sous l'action de la force F extérieure non seulement tout le travail de la force extérieure mais aussi une partie de l'énergie potentielle U de l'électron se transforment en énergie cinétique de sorte que $\Delta E_{cin} = A + \Delta U$. L'énergie cinétique E_{cin} et la vitesse v d'un tel électron croissent plus vite que celles de l'électron libre. Tout se passe comme si cet électron était devenu plus léger que l'électron libre, en possédant une masse effective $m_{eff} < m_0$.

Si la totalité du travail effectué par la force extérieure se transforme en énergie potentielle U de l'électron, c'est-à-dire que $A = \Delta U$, l'énergie cinétique et la vitesse de mouvement de l'électron ne subissent aucun accroissement — l'électron se comporte comme une particule de masse effective infiniment grande.

Enfin, si lors du mouvement de l'électron non seulement tout le travail de la force extérieure F mais aussi une partie de l'énergie cinétique E_{cin} que possède l'électron se transforment en énergie potentielle, de sorte que $\Delta U = A + \Delta E_{cin}$, la vitesse d'un tel électron diminue lors du mouvement dans le cristal, l'électron se ralentit, en se comportant comme une particule de *masse effective négative*. C'est précisément de cette façon que se comportent des électrons situés près du sommet de la bande d'énergie.

Près du niveau inférieur de la bande (au voisinage de $k = 0$), tant que l'énergie $E(k)$ de l'électron croît, lorsque k augmente, à

peu près proportionnellement à k^2, la vitesse du mouvement de translation de l'électron $v \sim dE/dk$ augmente proportionnellement à k, l'accélération du mouvement est positive et la masse effective $m_{eff} \sim (d^2E/dk^2)^{-1}$ conserve sa valeur positive constante m_n. Au point d'inflexion A de la courbe $E(k)$ la dérivée seconde $d^2E/dk^2 = 0$ et la dérivée première dE/dk passe par son maximum. Aussi, lorsqu'on s'approche de ce point, $m_{eff} \to \infty$ et $v \to v_{max}$. Au-delà du point d'inflexion la dérivée dE/dk commence à décroître de sorte que la vitesse v décroît elle aussi ; l'accélération devient ainsi négative, ce qui est équivalent, le sens d'action de la force extérieure F étant conservé, au changement du signe positif de la masse effective en signe négatif. Près du sommet de la bande d'énergie $E(k)$ devient de nouveau une fonction quadratique de k et la masse effective atteint sa valeur négative constante m_n'.

§ 4.6. Remplissage des bandes en électrons et division des corps en métaux, diélectriques et semiconducteurs

Nombre d'états dans une bande d'énergie. Dans chaque bande d'énergie les états des électrons diffèrent l'un de l'autre par la valeur de **k**. Dans le cas d'un cristal de dimensions limitées seules des valeurs discrètes déterminées de **k** satisfont aux conditions aux limites relatives à l'équation de Schrödinger. Pour les mêmes raisons que dans le cas des ondes élastiques (v. § 3.1), ces conditions aux limites sont généralement choisies périodiques et alors les valeurs permises des vecteurs d'onde des électrons sont les mêmes que celles des vecteurs d'onde des vibrations du réseau cristallin (v. fig. 3.5). Le nombre total de différentes valeurs de **k** dans la première zone de Brillouin est alors égal au nombre N d'atomes du cristal. Mais à chaque valeur de **k** de l'électron correspondent deux états et non pas un seul, qui diffèrent par la valeur du spin de l'électron. Ainsi une bande d'énergie contient $2N$ états de l'électron. Suivant le principe de Pauli une telle bande peut contenir au maximum $2N$ électrons. Notons que tant que N atomes se trouvaient loin l'un de l'autre, le niveau d'énergie correspondant de chaque atome pouvait contenir deux électrons de spins opposés de sorte que si ce niveau dans l'atome était occupé, la bande d'énergie formée depuis ce niveau est maintenant occupée dans le cristal.

Lien entre le remplissage des bandes et les propriétés électriques des solides. Le nombre d'électrons contenus dans un solide étant limité, seules quelques bandes d'énergie les plus basses sont entièrement remplies.

Suivant le caractère de remplissage des bandes par les électrons, tous les corps peuvent être subdivisés en deux grands groupes.

Le *premier groupe* réunit les corps dans lesquels au-dessus des

bandes entièrement remplies se situe une bande qui n'est remplie que de façon incomplète (fig. 4.8, *a*). Une telle bande apparaît dans le cas où le niveau atomique à partir duquel elle est formée, est rempli dans l'atome incomplètement, comme c'est le cas des métaux alcalins par exemple.

Une bande partiellement remplie peut également être créée par la superposition des bandes remplies et des bandes vides ou partielle-

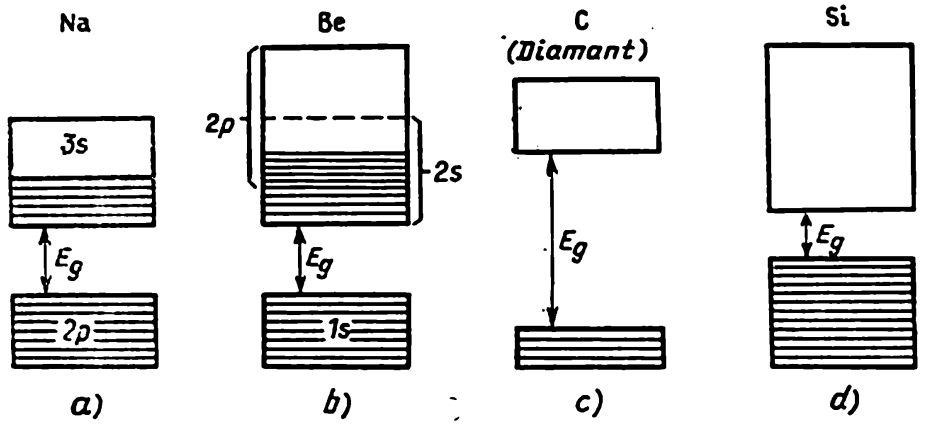

Fig. 4.8. Remplissage des bandes par les électrons dans les métaux (*a*, *b*), les diélectriques (*c*) et les semiconducteurs (*d*).

ment remplies comme cela a lieu pour le béryllium et les éléments alcalino-terreux (fig. 4.8, *b*).

La présence d'une bande remplie de façon incomplète est propre aux métaux.

Dans le *deuxième groupe* sont rangés des corps dans lesquels au-dessus des bandes complètement remplies se situent des bandes vides (fig. 4.8, *c*, *d*). Un exemple typique de tels corps sont les éléments chimiques du groupe IV de la classification périodique de Mendéléev — le carbone et les variétés du diamant, le silicium, le germanium et l'étain gris possédant la structure du diamant. Au même groupe de corps appartiennent de nombreux composés chimiques — les oxydes de métaux, les nitrures, les carbures, les halogénures de métaux alcalins, etc.

Suivant la théorie des bandes des solides, les électrons des bandes d'énergie périphériques possèdent pratiquement une même liberté de mouvement dans tous les corps qu'ils soient métaux ou diélectriques. Or, les propriétés électriques de ces corps et notamment la conductivité électrique diffèrent très fortement, de plusieurs ordres de grandeur: pour les métaux $\sigma \approx 10^7$ ohms$^{-1}\cdot$m^{-1}, pour les diélectriques de bonne qualité $\sigma < 10^{-11}$ ohms$^{-1}\cdot$m^{-1}.

Ainsi, l'existence d'électrons libres capables de se déplacer à travers le cristal n'est qu'une condition nécessaire mais non suffisante pour que les corps deviennent conducteurs de l'électricité.

Pour pouvoir établir la condition suffisante, elle aussi, de conductibilité électrique, considérons le comportement dans un champ extérieur des électrons situés dans des bandes d'énergie partiellement remplies et entièrement remplies du cristal.

Produisons dans un cristal un champ extérieur $\mathscr{E}$. Ce champ agit sur chaque électron avec une force $F = -e\mathscr{E}$ qui tend à troubler la symétrie de la répartition des électrons suivant les vitesses: les électrons qui se déplacent dans le sens opposé à cette force sont retardés, tandis que les électrons se déplaçant dans le sens de la force sont accélérés. L'accélération ou la décélération de l'électron est inévitablement liée à une variation de son énergie, ce qui signifie le passage de l'électron dans un nouvel état quantique ayant une énergie plus grande ou moins grande. *Il est évident que de telles transitions ne sont possibles que dans le cas où les bandes d'énergie auxquelles appartiennent des électrons donnés contiennent des états non occupés, c'est-à-dire si la bande n'est remplie que de façon incomplète.* Dans un tel cas un champ électrique même de faible intensité est capable de communiquer aux électrons une impulsion additionnelle suffisante pour les faire passer aux niveaux libres les plus proches. Le corps devient le siège d'un mouvement préférentiel des électrons dans le sens opposé au champ, ce qui constitue le passage d'un courant électrique. De tels corps doivent être de bons conducteurs de l'électricité, ce qui a lieu en réalité.

Imaginons maintenant que la plus élevée des bandes du cristal occupées par les électrons est entièrement remplie. Une telle bande est dite *bande de valence*. Supposons qu'elle est séparée de la bande suivante, libre, par une large fente énergétique E_g (fig. 4.8, *c*). Cet intervalle d'énergies interdites dans le cristal porte le nom de *bande interdite*. La bande libre située dans de tels corps au-dessus de la bande interdite est dite *bande de conduction* de même que la bande d'énergie remplie partiellement dans un métal qui assure une conductibilité électrique élevée de ce dernier. Le champ extérieur appliqué à un tel cristal n'est pas en mesure de changer le caractère de mouvement des électrons dans la bande de valence parce qu'il est incapable de faire passer les électrons dans la bande libre plus élevée. Quant à l'intérieur de la bande de valence elle-même, qui ne contient aucun niveau libre, le champ peut ici seulement provoquer une permutation des électrons, ce qui ne trouble pas la symétrie de la répartition suivant les vitesses. C'est pourquoi dans de tels corps le champ électrique extérieur n'est pas capable de faire apparaître un mouvement coordonné des électrons, c'est-à-dire un courant électrique, de sorte que ces corps doivent posséder une conductivité électrique pratiquement nulle.

Ainsi, *la condition suffisante d'apparition d'une conductivité élevée dans les corps est la présence dans leur spectre énergétique de bandes d'énergie qui ne sont remplies par les électrons que de façon*

incomplète, comme c'est le cas des métaux typiques (fig. 4.8, *a*, *b*). L'absence de telles bandes dans le spectre énergétique des corps solides du deuxième groupe les rend *non conducteurs* malgré le fait qu'ils contiennent des électrons se déplaçant à travers tout le cristal.

Suivant la largeur de la bande interdite on convient de diviser les corps du deuxième groupe en *diélectriques* et en *semiconducteurs*. Les diélectriques sont représentés par des corps dont la bande interdite est relativement large. Pour des diélectriques typiques $E_g > 3$ eV. C'est ainsi que pour le diamant $E_g = 5{,}2$ eV, pour le nitrure de bore $E_g = 4{,}6$ eV ; pour Al_2O_3 $E_g = 7$ eV, etc.

Dans la catégorie de semiconducteurs se rangent des corps possédant une bande interdite relativement étroite (fig. 4.8, *d*). Pour des semiconducteurs typiques $E_g \lesssim 1$ eV. C'est ainsi que pour le germanium $E_g = 0{,}66$ eV ; pour le silicium $E_g = 1{,}08$ eV ; pour l'antimoniure d'indium $E_g = 0{,}17$ eV ; pour l'arséniure de gallium $E_g = 1{,}43$ eV, etc.

Examinons de plus près ce groupe de corps.

§ 4.7. Semiconducteurs intrinsèques. Notions sur les trous

Semiconducteurs intrinsèques. Les semiconducteurs chimiquement purs sont appelés *semiconducteurs intrinsèques*. Ce sont certains éléments chimiquement purs (germanium, silicium, sélénium, tellure et autres) et de nombreux composés chimiques tels que par exemple l'arséniure de gallium GaAs, l'arséniure d'indium InAs, l'antimoniure d'indium InSb, le carbure de silicium SiC, etc.

Le schéma simplifié de la structure de bande d'un semiconducteur intrinsèque est représenté sur la figure 4.9, *a*. A $T = 0$ K sa bande de valence est complètement remplie, la bande de conduction située au-dessus de la bande de valence à une distance E_g est vide. Aussi, à $T = 0$ K un semiconducteur intrinsèque possède, de même qu'un diélectrique, une conductibilité électrique nulle.

Toutefois, lorsque la température s'élève, une partie des électrons de la bande de valence acquièrent, grâce à l'excitation thermique, une énergie suffisante pour franchir la bande interdite et passer dans la bande de conduction (fig. 4.9, *b*). Il en résulte l'apparition d'électrons libres dans la bande de conduction et de niveaux libres dans la bande de valence auxquels peuvent passer les électrons de cette bande. Si un tel cristal est soumis à l'action d'un champ électrique extérieur, il devient le siège d'un mouvement ordonné de l'ensemble des électrons de la bande de conduction et de la bande de valence qui constitue un courant électrique. Le cristal devient *conducteur de l'électricité*.

Les électrons reçoivent l'énergie E_g, nécessaire pour passer de la bande de valence dans la bande de conduction, de la part des atomes

vibrants du cristal et du rayonnement en équilibre thermique. Comme l'énergie moyenne des vibrations de l'atome ($\sim kT$) est d'ordinaire très inférieure à l'énergie E_g, seule une très petite partie des atomes peuvent exciter les électrons de la bande de valence. Dans le rayonnement en équilibre thermique lui aussi la fraction des photons ayant une énergie $\hbar\omega > E_g$ est petite. Lorsque la température s'élève, ces fractions augmentent de sorte que plus étroite est la bande interdite et plus élevée est la température du cristal, plus grand est le nombre des électrons qui passent dans la bande de conduction et plus haute

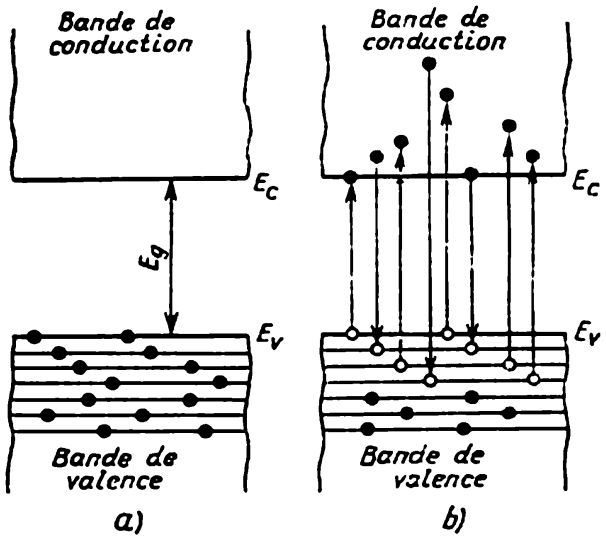

Fig. 4.9. Remplissage des bandes par les électrons dans un semiconducteur à $T = 0$ K (*a*) et transitions des électrons dans la bande de conduction en cas d'élévation de la température (*b*).

devient la conductivité électrique du cristal. C'est ainsi que dans le germanium dont $E_g = 0{,}66$ eV, la concentration du gaz électronique dans la bande de conduction atteint, même à la température ordinaire, une valeur $n_i \approx 10^{19}$ m^{-3} et la résistivité n'est égale qu'à $\rho_i \approx 0{,}48$ ohm·m. En même temps, pour le phosphure de gallium, dont $E_g = 2{,}24$ eV, on a à la température ordinaire $n_i \approx 10^{-6}$ m^{-3} et $\rho_i \approx 2 \cdot 10^{13}$ ohms·m. Mais à $T = 900$ K, la concentration du gaz électronique dans le phosphure de gallium augmente de plusieurs ordres de grandeur et la résistivité électrique devient de l'ordre de grandeur de la résistivité du germanium à la température ordinaire.

De ce qui précède on peut tirer deux conclusions importantes suivantes :

1. La conductibilité des semiconducteurs est une conductibilité *induite* : elle apparaît sous l'action d'un facteur extérieur capable de communiquer aux électrons de la bande de valence une énergie suffisante pour les faire passer dans la bande de conduction. Ces

facteurs peuvent être l'échauffement des semiconducteurs, le rayonnement lumineux et le rayonnement ionisant.

2. La division des corps en semiconducteurs et en diélectriques est pour une large part arbitraire. Le phosphure de gallium, qui est diélectrique à la température ordinaire, acquiert une conductibilité notable aux températures plus élevées et peut donc être considéré comme un semiconducteur.

Notions sur les trous. Examinons maintenant plus en détail le comportement des électrons dans la bande de valence où il apparaît des niveaux libres du fait que certains électrons sont passés dans la bande de conduction (fig. 4.9, *b*).

Les électrons de la bande de valence peuvent maintenant passer, sous l'action du champ extérieur, aux niveaux libres et créer dans le cristal un courant électrique.

La contribution à la densité de courant apportée par un électron se déplaçant à la vitesse $\mathbf{v}_j$ dans un conducteur de volume unité a pour valeur

$$\mathbf{i}_j = -e\mathbf{v}_j.$$

Tous les électrons de la bande de valence produisent une densité de courant

$$\mathbf{i} = -e\sum_j \mathbf{v}_j,$$

où la sommation est étendue à tous les états occupés par les électrons.

Les courbes de dispersion sont symétriques dans toute bande d'énergie, c'est-à-dire que les états électroniques permis peuvent être groupés en paires dont les sens des vitesses sont opposés. Une de telles paires d'états de vitesses v_s et $-v_s$ est montrée sur la figure 4.10, *a*. Ainsi, le courant engendré par les électrons de la bande complètement remplie est nul.

Imaginons maintenant que dans la bande de valence tous les états sont occupés excepté un seul caractérisé par la vitesse $\mathbf{v}_s$. La densité totale du courant dans une telle bande a pour expression

$$\mathbf{i} = -e\sum_{j\neq s} \mathbf{v}_j = -e\sum_j \mathbf{v}_j + e\mathbf{v}_s.$$

Comme le premier terme du second membre de cette égalité est nul, on a

$$\mathbf{i} = e\mathbf{v}_s.$$

Ainsi, *le courant total de tous les électrons de la bande de valence possédant un état vacant est équivalent au courant produit par le mouvement d'une particule de charge positive $+e$ placée dans cet état.* Une telle particule fictive a reçu le nom de *trou*.

Regardons maintenant comment se déplace le trou sous l'effet d'un champ électrique extérieur. Notons tout d'abord qu'en état

d'équilibre thermique les électrons tendent à occuper les niveaux d'énergie les plus bas de sorte que le trou se trouve placé au sommet de la bande de valence où sa vitesse est nulle (fig. 4.10, *a*). Appliquons maintenant le champ électrique extérieur dans le sens de l'axe des x. Alors la force s'exerçant sur la particule positive — le trou,

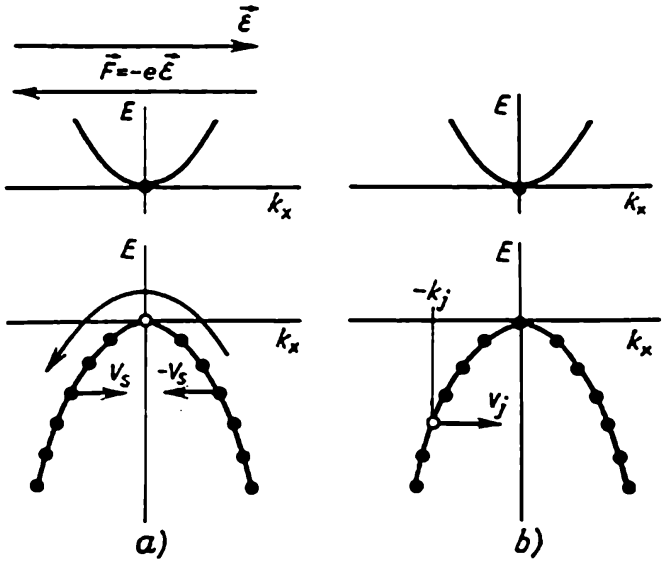

Fig. 4.10. Explication de la notion de trou.

sera dirigée elle aussi suivant l'axe des x et, si la masse effective du trou est positive, le trou devrait se déplacer, au bout d'un certain temps, avec une vitesse dirigée suivant l'axe des x. Regardons dans quel sens se déplace le trou en réalité.

Le champ extérieur agit sur les électrons dans le sens des x négatifs, en tendant à faire varier leur vecteur d'onde dans ce sens. Ceci étant, les électrons se déplacent, tous à la fois, sur la courbe de dispersion dans le sens indiqué par une flèche sur la figure 4.10, *a*, de sorte qu'au bout d'un certain temps le trou se trouve dans l'état $-k_j$ (fig. 4.10, *b*) où sa vitesse v_j est réellement dirigée suivant l'axe des x, c'est-à-dire dans le sens de la force qui lui est appliquée. Par conséquent, *on doit attribuer au trou une masse effective positive numériquement égale à la masse effective négative de l'électron se trouvant près du sommet de la bande de valence.*

§ 4.8. Surfaces de Fermi des métaux

Le niveau le plus élevé parmi les niveaux occupés (à $T = 0$ K) de la bande de conduction d'un métal est appelé *niveau de Fermi* et son énergie est dite *énergie de Fermi* μ. A l'énergie de Fermi correspondent des valeurs déterminées des vecteurs d'onde $\mathbf{k}$ des électrons données par la relation de dispersion $E = E(\mathbf{k})$. De telles valeurs

sont concentrées sur une certaine surface dans l'espace des vecteurs d'onde (espace des **k**) que l'on appelle *surface de Fermi*. La forme de cette surface a une influence considérable sur de nombreuses propriétés électriques, magnétiques et autres des métaux. La forme de cette surface est la plus simple pour des métaux monovalents. Revenons à titre d'exemple à l'examen de la bande de conduction du sodium (v. § 4.1). Cette bande est formée à partir des niveaux d'énergie $3s$ des atomes de Na. Elle contient $2N$ états, où N est le nombre

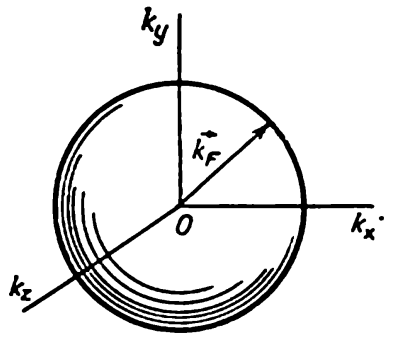

Fig. 4.11. Surface de Fermi sphérique d'un métal monovalent.

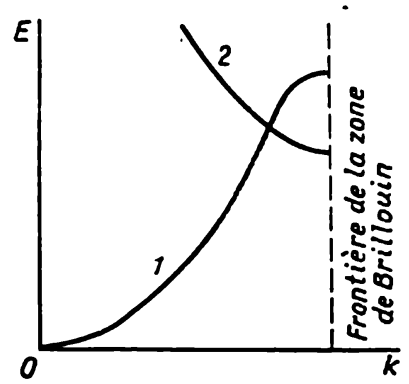

Fig. 4.12. Recouvrement des bandes d'énergie dans les métaux.

d'atomes dans le cristal (v. § 4.6), et est occupée par N électrons — un électron de chaque atome. Dans un tel cas cette bande est occupée à moitié. Dans les métaux, l'énergie d'un électron ne dépend pas de l'orientation de son vecteur d'onde et est liée de façon approchée (dans l'approximation de l'électron libre) au module de ce vecteur par la relation (4.13). Alors pour les électrons de Fermi, c'est-à-dire les électrons ayant l'énergie de Fermi μ, le module du vecteur d'onde a pour valeur

$$k_F = \sqrt{2m_e\mu}/\hbar.$$

La surface de Fermi est dans ce cas une sphère de rayon k_F dans l'espace des **k** (fig. 4.11).

La situation est loin d'être si simple dans le cas des cristaux d'éléments bivalents, par exemple de béryllium (Be). Il semble à première vue que de tels cristaux doivent être diélectriques parce que leur nombre d'électrons de valence ($2N$) est précisément égal au nombre d'états dans la bande d'énergie correspondante. Mais la bande d'électrons de valence des éléments alcalino-terreux se recouvre en partie avec la bande libre suivante (fig. 4.8, *b*). Ce recouvrement se produit près des frontières de la zone de Brillouin dans certaines directions dans l'espace des **k** (courbes de dispersion *1* et *2* de la figure 4.12). Dans un tel cas une partie des électrons de la bande de valence se trouvent situés dans la bande libre suivante. Il se forme

ainsi dans la bande de valence un certain nombre d'états libres — de trous et, dans la bande de conduction, un certain nombre d'électrons. A la différence des semiconducteurs, l'apparition d'électrons et de trous dans de tels métaux n'est pas due à l'excitation thermique mais, de même que dans les semiconducteurs intrinsèques, le courant dans ces métaux est transporté par des particules de deux signes — par des électrons chargés négativement et par des trous chargés positivement.

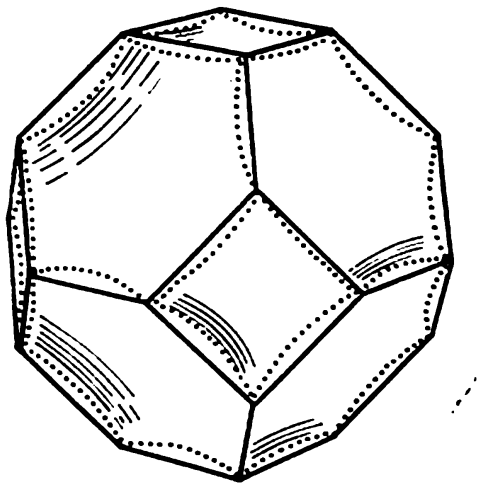

Fig. 4.13. Surface de Fermi (contour en pointillé) de la deuxième des bandes qui se recouvrent dans l'Al (la première bande est entièrement occupée).

Des phénomènes analogues sont également observés dans des métaux de valence plus élevée où les électrons de valence peuvent occuper totalement ou partiellement plusieurs bandes d'énergie. Pour les bandes d'énergie qui se chevauchent les surfaces de Fermi peuvent avoir une forme bien compliquée. La figure 4.13 montre à titre d'exemple la surface de Fermi pour l'aluminium (qui est loin d'être la plus complexe). La zone de Brillouin n'est pas dans ce cas un cube parce que le cristal d'aluminium est un cristal à faces centrées et non pas cubique simple.

§ 4.9. Niveaux localisés dans la bande interdite

Les semiconducteurs contiennent toujours, quel que soit leur degré de pureté, des atomes d'impuretés qui créent leurs propres niveaux énergétiques qu'on appelle *niveaux d'impuretés*. Ces niveaux peuvent se situer aussi bien dans les bandes permises que dans les bandes interdites à différentes distances du sommet de la bande de valence et du niveau inférieur de la bande de conduction. Dans certains cas, les impuretés sont introduites intentionnellement pour conférer au semiconducteur des propriétés désirées.

Examinons les principaux types de niveaux d'impuretés.

Niveaux donneurs. Supposons que dans un cristal de germanium une partie d'atomes de germanium sont remplacés par des atomes d'arsenic pentavalent. Le germanium possède le même réseau que le diamant dans lequel chaque atome est entouré par quatre voisins les plus proches qui lui sont liés par des liaisons de valence (fig. 4.14, *a*). De chacun des quatre voisins l'atome de germanium reçoit sur ses orbitales libres un électron et, à son tour, cède ses quatre

électrons de valence aux atomes voisins. C'est ainsi que toutes les huit orbitales de la couche de valence du germanium se trouvent occupées. Quant à l'atome d'arsenic introduit dans le réseau cristallin de germanium, il établit lui aussi des liaisons avec quatre atomes de germanium les plus proches. Mais au cinquième électron de la couche de valence de l'arsenic il ne reste plus de place sur cette couche

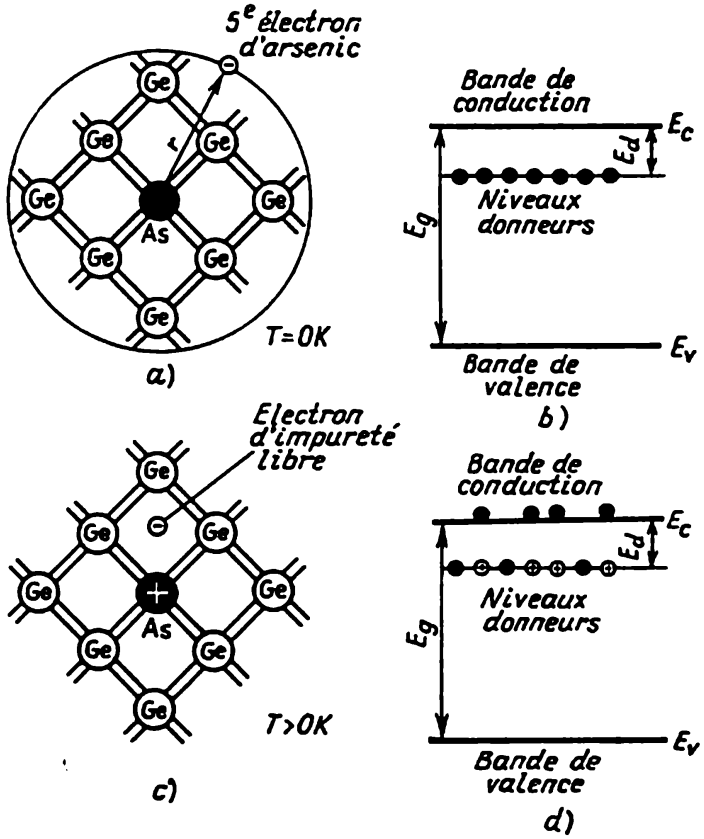

Fig. 4.14. Atome donneur dans un semiconducteur à $T = 0$ K (*a*) et à $T > 0$ K (*c*); remplissage du niveau donneur par les électrons à $T = 0$ K (*b*) et à $T > 0$ K (*d*).

de sorte qu'il doit passer sur une orbite plus éloignée. Toutefois entre cet électron et l'ion arsenic se trouvent maintenant des nuages électroniques facilement polarisables de liaisons de valence (la permittivité relative du germanium assurée par ces nuages $\varepsilon = 16$). L'attraction de l'électron par l'ion arsenic devient ε fois plus faible et l'orbite de l'électron augmente encore.

A une grande distance le champ engendré par l'ion arsenic coïncide pratiquement avec le champ d'une charge ponctuelle $+e$. Ceci étant, l'équation de Schrödinger pour l'électron considéré dans le champ de l'ion a la même forme que pour l'électron dans l'atome d'hydrogène, à cette différence près qu'au lieu de la permittivité ε_0 elle doit faire intervenir la quantité $\varepsilon\varepsilon_0$ et au lieu de la masse réelle de l'électron sa masse effective dans la bande de conduction. En effet,

l'électron considéré se déplace à travers le cristal en dehors du système de liaisons de valence, c'est-à-dire comme un électron de la bande de conduction.

Dans un tel cas l'énergie E_d de liaison de l'électron avec l'ion est donnée par une formule analogue à (4.2):

$$E_d = \frac{e^4 m_n}{8(\varepsilon\varepsilon_0)^2 h^2}. \tag{4.20}$$

En tenant compte que dans le germanium $m_n = 0{,}22m_0$ et $\varepsilon = 16$, on obtient $E_d \approx 0{,}01$ eV. Si l'on communique à l'électron une telle énergie, il se détachera de l'atome et sera capable de se déplacer librement dans le réseau cristallin du germanium en devenant ainsi un électron de conduction (fig. 4.14, *c*).

Dans le langage de la théorie des bandes ce processus peut être décrit de la façon suivante. Entre la bande de valence remplie et la bande de conduction libre on voit se situer les niveaux énergétiques des cinquièmes électrons de l'atome d'arsenic (fig. 4.14, *b*). Ces niveaux se disposent immédiatement près du niveau inférieur de la bande de conduction à une distance $E_d \approx 0{,}01$ eV. Lorsque les électrons occupant de tels niveaux d'impuretés reçoivent une énergie E_d, ils passent dans la bande de conduction (fig. 4.14, *d*). Les charges positives qui en résultent se localisent sur les atomes d'arsenic immobiles et ne participent pas à la conduction électrique.

Les impuretés qui constituent la source d'électrons de conduction sont dites *donatrices* ou *donneurs* et les niveaux énergétiques de ces impuretés sont appelés *niveaux donneurs*. Les semiconducteurs contenant une impureté donatrice sont appelés *semiconducteurs par excès d'électrons* ou *semiconducteurs du type n*.

Niveaux accepteurs. Supposons maintenant que dans le réseau de germanium une partie des atomes de germanium sont remplacés par des atomes d'indium trivalent (fig. 4.15, *a*). Pour établir les liaisons avec les quatre voisins les plus proches dans l'atome d'indium manque un électron. C'est pourquoi, près de cet atome une des liaisons n'est pas complète (fig. 4.15, *a*). Or, en réalité, la liaison non complète ne doit pas rester obligatoirement directement près de l'atome d'indium. Le fait est que les électrons de la bande de valence ne sont nullement fixés chacun dans une liaison déterminée mais peuvent se déplacer dans le système de liaisons de valence. Ainsi, pour l'instant il est seulement clair que dans le système de liaisons de valence il existe une vacance. Mais, comme il a été montré au § 4.7, le comportement de tous les autres électrons est dans ce cas équivalent au comportement d'une quasi-particule unique — d'un trou. Ainsi, l'introduction d'un atome d'un élément appartenant au groupe III du tableau périodique de Mendéléev dans le germanium a eu pour résultat l'apparition d'un trou dans la bande de valence. Toutefois ce trou doit être lié à l'atome d'impureté. En effet, lorsque toutes les

liaisons près de l'atome d'indium sont complètes, cet atome devient un ion chargé négativement (fig. 4.15, *c*) et attire le trou chargé positivement. Vu que le couplage entre le trou et l'ion s'affaiblit du fait de la polarisation du cristal, cette liaison est suffisamment faible. Il est facile de voir qu'il existe une analogie complète entre l'examen d'un électron tournant autour d'un ion positif dans le semiconducteur et l'examen d'un trou tournant autour d'un ion négatif.

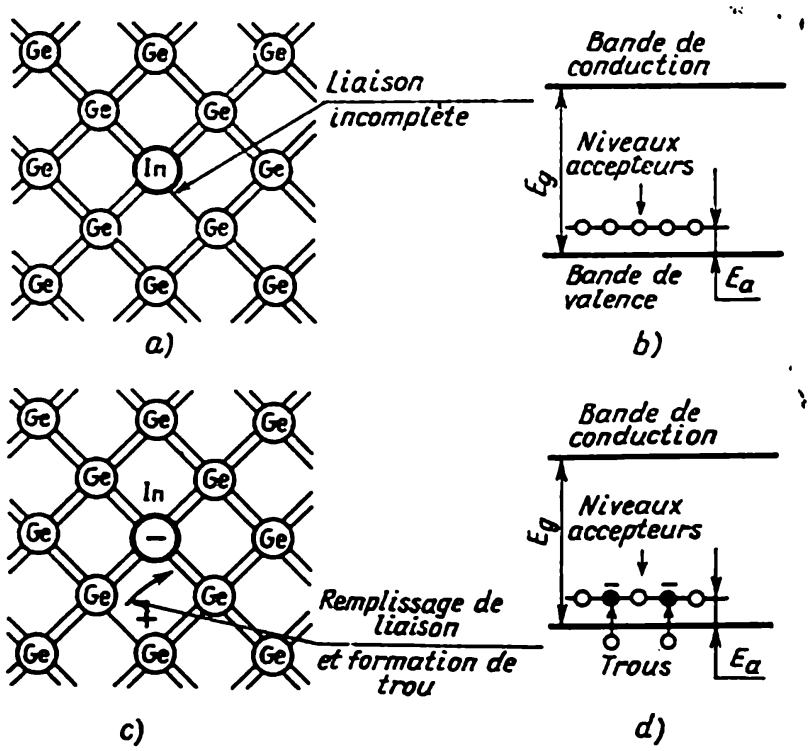

Fig. 4.15. Atome accepteur dans un semiconducteur à $T = 0$ K (*a*) et à $T > 0$ K (*c*) ; remplissage du niveau accepteur à $T = 0$ K (*b*) et à $T > 0$ K (*d*).

L'énergie de liaison entre le trou et l'ion peut être obtenue si dans la formule (4.20) on change m_n et m_p. Si l'on fournit à un trou lié une énergie égale à l'énergie de sa liaison $E_a \approx 0{,}01$ eV, il se détache de l'ion d'indium et devient libre. Ce processus est illustré par la figure 4.15, *b* qui montre la structure des bandes d'énergie du germanium contenant une impureté d'indium. Les niveaux non occupés des atomes d'indium se situent directement près du sommet de la bande de valence à une distance $E_a \approx 0{,}01$ eV. Le fait que ces niveaux se situent au voisinage immédiat de la bande de valence a pour résultat que les électrons passent de la bande de valence aux niveaux d'impuretés même à des températures relativement peu élevées (fig. 4.15, *d*). En entrant en liaison avec les atomes d'indium, ils perdent la faculté de se déplacer dans le réseau de germanium et ne participent pas à la conduction. Les porteurs de charge ne sont cons-

tués que par les trous qui prennent naissance dans la bande de valence.

Les impuretés qui captent les électrons de la bande de valence d'un semiconducteur sont appelées *impuretés acceptrices* ou *du type p* et leurs niveaux d'énergie sont appelés *niveaux accepteurs*. Les semiconducteurs contenant de telles impuretés sont dits *semiconducteurs par défaut d'électrons* ou *semiconducteurs du type p*.

Les atomes d'impuretés des autres groupes du tableau de Mendéléev forment des niveaux donneurs situés loin de la bande de conduction et des niveaux accepteurs situés loin de la bande de valence. De ce fait ils n'influent pratiquement pas sur la conductibilité électrique des semiconducteurs. En revanche, ils ont une grande influence sur les processus de génération et de recombinaison d'électrons et de trous (v. chap. 8). Ce que nous venons de dire pour les atomes en position de substitution s'applique intégralement aux atomes situés non pas aux nœuds du réseau cristallin mais aux interstrices (*atomes en position interstitielle*). Les vacances et d'autres imperfections du réseau cristallin engendrent elles aussi des niveaux d'énergie dans la bande interdite du semiconducteur qui peuvent être aussi bien donneurs qu'accepteurs et se situent en règle générale à une profondeur suffisamment grande dans la bande interdite.

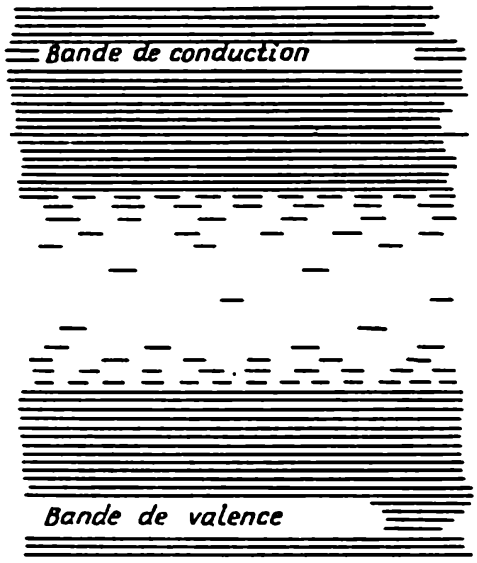

Fig. 4.16. Diagramme énergétique d'un semiconducteur amorphe.

Les fonctions d'onde des électrons (ou des trous) occupant les niveaux d'impuretés sont localisées au voisinage des atomes ou des défauts correspondants de la structure, ce qui explique le qualificatif de *localisé* utilisé pour désigner ces niveaux. Toutefois, dans le cas où la concentration d'atomes d'impureté est forte, les électrons qui leur sont liés deviennent capables de se déplacer d'un atome d'impureté à un autre et les niveaux localisés sont si serrés qu'ils ont l'aspect d'une bande d'impureté.

Structure de bande des semiconducteurs amorphes. Des substances amorphes elles aussi peuvent posséder les propriétés semiconductrices de même que les cristaux. Dans les solides amorphes il n'existe pas d'ordre à longue distance dans l'arrangement des atomes propre aux cristaux, mais le caractère des liaisons et les distances entre les atomes

voisins restent les mêmes que dans les cristaux. Autrement dit, les facteurs qui conduisent à la mise en commun des électrons de valence et à la délocalisation de leurs fonctions d'onde se conservent. Aussi, de même que dans les cristaux, le spectre énergétique des électrons mis en commun a-t-il un caractère de bande. Toutefois dans la bande interdite des semiconducteurs amorphes il existe un grand nombre de niveaux localisés. L'origine de tels niveaux est facile à comprendre si l'on se rappelle que dans les semiconducteurs cristallins eux-mêmes les défauts de structure produisent des niveaux localisés dans la bande interdite et une substance amorphe peut être considérée comme un cristal dont le nombre de défauts est aussi grand que possible. Le spectre énergétique d'un semiconducteur amorphe est représenté schématiquement sur la figure 4.16 où les lignes continues désignent les états des électrons mis en commun et les traits figurent les états localisés.

§ 4.10. Quantification de l'énergie de l'électron dans un champ magnétique. Niveaux de Landau

Si l'électron comporte une composante de vitesse normale au vecteur induction d'un champ magnétique, il subit l'action d'une force dite de Lorentz qui incurve sa trajectoire. Dans un plan normal au champ, l'électron doit être animé d'un mouvement suivant une circonférence dont le rayon est proportionnel à la vitesse de l'électron et inversement proportionnel à l'induction du champ magnétique. La fréquence cyclique de révolution d'un tel électron suivant la circonférence (dite *fréquence cyclotron* ω_c) a pour valeur

$$\omega_c = eB/m_e, \tag{4.21}$$

où B est l'induction de champ magnétique ; m_e la masse de l'électron.

L'énergie cinétique du mouvement de l'électron dans un plan perpendiculaire à **B** se détermine par l'égalité

$$E_\perp = {}^1/_2\, m_e r^2 \omega_c^2, \tag{4.22}$$

où r est le rayon de l'orbite de l'électron.

Puisque du point de vue classique le rayon de l'orbite électronique peut prendre une suite continue de valeurs, l'énergie $E_\perp$ peut elle aussi prendre toute valeur. L'étude quantique conduit à la conclusion que l'énergie $E_\perp$ ne peut prendre qu'une suite discontinue de valeurs :

$$E_\perp = (n + {}^1/_2)\, \hbar\omega_c, \tag{4.23}$$

où $n = 0, 1, 2, \ldots$ (la formule analogue (3.23) déterminait l'énergie d'une vibration normale du réseau).

L'électron se déplaçant le long du champ **B** n'étant pas soumis à l'action de la force de Lorentz, l'énergie $E_\parallel$ d'un tel mouvement se détermine comme précédemment [v. (4.13)] par la relation

$$E_{\parallel} = \frac{\hbar^2 k_{\parallel}^2}{2m_e}, \tag{4.24}$$

où $k_{\parallel}$ est la composante du vecteur d'onde de l'électron dans le sens du champ magnétique. Ainsi, l'énergie cinétique totale de l'électron dans un champ magnétique a pour expression

$$E = E_{\perp} + E_{\parallel} = \left(n + \frac{1}{2}\right) \hbar\omega_c + \frac{\hbar^2 k_{\parallel}^2}{2m_e}. \tag{4.25}$$

Le caractère du mouvement d'un électron de la bande de conduction du cristal dans un champ magnétique peut différer de façon substantielle de celui que nous venons de décrire parce que dans les cristaux réels les électrons diffusent constamment, changent le sens de leur mouvement par suite de l'interaction avec tous les défauts différents possibles du réseau. Toutefois, dans un champ magnétique suffisamment intense, la longueur de la circonférence décrite par l'électron, dans la représentation classique, ou la localisation de sa fonction d'onde dans une direction normale à **B**, en théorie quantique, peuvent devenir beaucoup plus petites que la longueur de libre parcours moyen de sorte que les représentations et les formules indiquées plus haut deviennent valables. Il suffit de remplacer seulement la masse réelle m_e de l'électron par sa masse effective m_n.

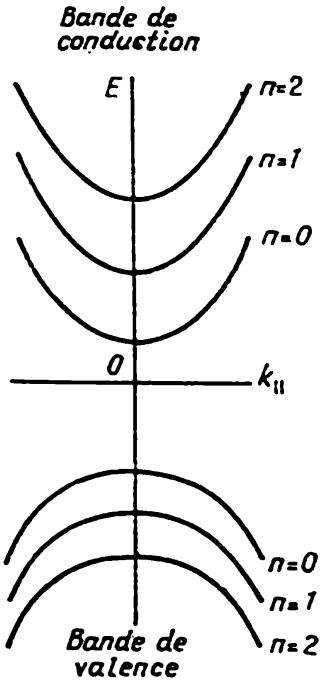

Fig. 4.17. Courbes de dispersion pour un semiconducteur placé dans un champ magnétique intense.

La figure 4.17 représente schématiquement la variation de l'énergie des électrons d'un cristal dans un champ magnétique fort en fonction du vecteur d'onde $k_{\parallel}$ suivant l'expression (4.25). Pour une valeur fixée de $k_{\parallel}$, par exemple pour $k_{\parallel} = 0$, on obtient pour chaque bande un système de niveaux d'énergie discrets équidistants que l'on appelle *niveaux de Landau*. La distance $\hbar\omega_c$ entre les niveaux de Landau n'est pas en réalité si grande comme le montre la figure 4.17. Dans des champs de $\sim$1 T elle est égale à $\sim 10^{-4}$ eV. Chacune des courbes de dispersion représentées sur la figure 4.17 se décompose encore en deux courbes : pour le spin de l'électron orienté suivant le champ et pour le spin orienté dans le sens opposé au champ magnétique (sur la figure 4.17 cette décomposition n'est pas représentée).

CHAPITRE 5

STATISTIQUE DES ÉLECTRONS ET DES TROUS DANS LES SEMICONDUCTEURS

Tout corps solide représente un système ou une *collectivité* constitué d'un énorme nombre de particules microscopiques. De tels systèmes obéissent à des lois spécifiques dites *statistiques* qui font l'objet d'étude de la physique statistique. Dans ce chapitre on obtiendra, sur la base de ces lois, des expressions très importantes pour des applications pratiques donnant les concentrations des porteurs de charges libres (électrons et trous) dans les semiconducteurs ainsi que leur répartition entre divers niveaux d'énergie, etc.

§ 5.1. Fonction de répartition dans la statistique de Fermi-Dirac

En mécanique classique, l'état de chaque particule constitutive de la collectivité est décrit par la donnée de ses trois coordonnées et de trois composantes de l'impulsion. Il semblerait qu'en établissant les équations du mouvement de ces particules et en les résolvant, on pourrait obtenir des renseignements complets sur le comportement du système et prédire son état à tout instant. Mais les calculs de ce genre sont non seulement extrêmement complexes, ils sont inutiles. La complexité du problème est illustrée par le fait que pour décrire le comportement des molécules de gaz contenues dans 1 m^3 sous des conditions normales il faudrait résoudre près de 10^{24} équations du mouvement liées entre elles compte tenu des conditions initiales, ce qui est pratiquement impossible à réaliser. Toutefois, si une telle résolution était effectuée, elle serait inutile parce que les propriétés du système arrivé à un état d'équilibre non seulement ne dépendent pas des valeurs initiales des coordonnées et des composantes de l'impulsion mais restent inchangées avec le temps bien que les coordonnées et les impulsions des particules varient sans cesse. Il en résulte que la collectivité considérée comme un tout est un système qualitativement différent des particules distinctes et que son comportement obéit à des lois autres que celles régissant les particules distinctes, à savoir aux lois statistiques. L'existence de telles lois est confirmée par des exemples suivants.

La vitesse d'une molécule de gaz distincte est une grandeur aléatoire dont la valeur est impossible à prédire. Malgré cela, dans un gaz constitué d'un grand nombre de telles molécules on observe une répartition bien déterminée de ces molécules suivant les vitesses:

une partie bien déterminée de molécules possèdent en moyenne une vitesse comprise entre 100 et 200 m/s, entre 400 et 500 m/s, etc.

La présence d'une molécule donnée dans un élément de volume isolé est un effet du hasard. Malgré cela on constate que la répartition des molécules suivant le volume obéit à une loi bien déterminée: des éléments de volume égaux contiennent en moyenne un même nombre de molécules.

La situation qui se produit ici est à peu près la même que dans le cas du jet d'une pièce de monnaie. La tombée de la pièce du côté face ou pile à chaque jet est un événement aléatoire. Néanmoins, pour un nombre suffisamment grand de jets, on constate une loi bien déterminée: en moyenne, la moitié de jets marque face et la moitié pile. Ce sont les lois de ce genre qu'on appelle lois statistiques.

La particularité principale des lois statistiques est leur caractère *probabiliste*. Elles permettent de prédire seulement la probabilité d'un tel ou tel événement, de réalisation d'un tel ou tel résultat.

C'est ainsi que dans l'exemple de lancement d'une pièce de monnaie la probabilité de tombée sur un côté ou sur un autre est égale à 1/2. Des épreuves individuelles peuvent et doivent présenter des écarts par rapport à cette valeur de la probabilité, ces écarts étant d'autant plus grands que le nombre de telles épreuves est plus petit. Dans cinq épreuves successives la pièce peut marquer face 3, 4 et même 5 fois. Toutefois, plus grand est le nombre d'épreuves, c'est-à-dire le nombre d'éléments constituant la collectivité, plus déterminées et précises deviennent les prédictions statistiques.

Les calculs montrent que pour un système comportant N particules, l'écart relatif de la grandeur physique observée M par rapport à sa valeur moyenne $\overline{M}$ est inversement proportionnel à $\sqrt{N}$:

$$\frac{\sqrt{\overline{(\Delta M)^2}}}{\overline{M}} \sim \frac{1}{\sqrt{N}}.$$

Lorsque N augmente, ce rapport tend vers zéro et $M \to \overline{M}$. Pour une valeur suffisamment grande de N on obtiendra $M \approx \overline{M}$.

Prenons, à titre d'exemple, pour M le nombre de molécules d'air se trouvant dans un volume donné. Dans les conditions normales, 1 m³ d'air contient en moyenne $2{,}7 \cdot 10^{25}$ molécules. L'écart relatif par rapport à ce nombre est en moyenne égal à $1/\sqrt{N} \approx 2 \cdot 10^{-13}$. Cet écart est si infime qu'en pratique on peut le négliger et considérer que la répartition des molécules en volume est uniforme.

Il convient pourtant de souligner que les écarts par rapport à la valeur moyenne non seulement peuvent mais doivent avoir lieu. De tels écarts sont appelés *fluctuations*.

Le nombre moyen de particules d'un système dont l'énergie est comprise entre E et $E + dE$ est déterminé par la *fonction de répartition statistique complète* $N(E)\,dE$.

La fonction de répartition complète peut être représentée sous la forme du produit du nombre d'états $g(E)\,dE$ contenus dans l'inter-

valle d'énergie dE par la probabilité de remplissage de ces états par les particules. Notons cette fonction $f(E)$. Alors il vient

$$N(E)\,\mathrm{d}E = f(E)\,g(E)\,\mathrm{d}E. \tag{5.1}$$

La fonction $f(E)$ est appelée tout simplement *fonction de répartition*. Comme il a été dit plus haut, elle exprime la densité moyenne de remplissage des états donnés par les particules.

Si, par exemple, on a en moyenne 10 particules par 100 états situés l'un près de l'autre, la probabilité de remplissage de ces états $f(E) = 0{,}1$. Vu qu'en moyenne on a 0,1 particule par chaque état, la fonction $f(E)$ peut s'interpréter comme le nombre moyen de particules se trouvant dans l'état donné.

Ainsi, le problème de la détermination de la fonction de répartition complète des particules suivant les états se ramène à la recherche de la fonction $g(E)\,\mathrm{d}E$ qui décrit la répartition des états entre divers niveaux d'énergie et de la fonction $f(E)$ qui détermine la densité moyenne de remplissage de ces états par les particules.

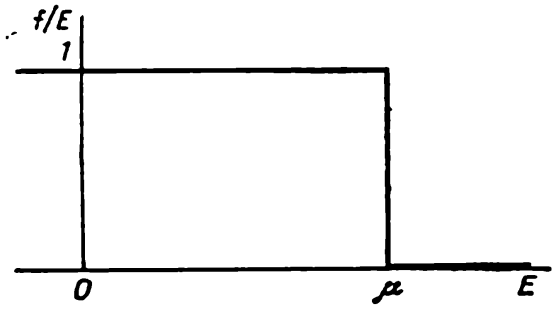

Fig. 5.1. Fonction de Fermi-Dirac à $T = 0$ K.

La *fonction de répartition pour les fermions*, et notamment pour les électrons, est de la forme

$$f(E) = (\mathrm{e}^{(E-\mu)/kT} + 1)^{-1}, \tag{5.2}$$

où T est la température thermodynamique et μ le potentiel chimique du système qui joue le rôle du paramètre de répartition et est appelé énergie de Fermi.

La fonction (5.2) porte le nom de *fonction de Fermi-Dirac*. Examinons sa forme d'abord dans le cas où $T \to 0$. Dans ce cas, $f(E) = 1$ pour toute énergie E inférieure à μ et $f(E) = 0$ pour $E > \mu$ (fig. 5.1). Ainsi, à $T = 0$ K, tous les états d'énergie $E < \mu$ sont occupés par les électrons et les états de $E > \mu$ sont libres. En d'autres termes, lorsque $T \to 0$ K, les électrons tendent à occuper les niveaux d'énergie les plus bas. Lorsque la température augmente, les électrons subissent une excitation thermique et passent à des niveaux d'énergie plus hauts de sorte que le caractère de leur répartition suivant les états change. Cependant à l'excitation thermique ne peuvent être soumis que des électrons situés dans une bande étroite $\sim kT$ disposée directement près du niveau de Fermi (sur la figure 5.2, *a* les états excités sont marqués par hachure et les sens des spins des électrons sont indiqués par des flèches). Les électrons occupant des niveaux plus profonds restent pratiquement à leurs places parce que l'énergie kT d'agitation thermique des atomes est insuffisante pour leur excitation (pour les faire passer au-delà du niveau de Fermi).

Par suite de l'excitation thermique une partie des électrons ayant une énergie inférieure à μ, passent à des niveaux d'énergie supérieure à μ si bien qu'il s'établit une nouvelle répartition des électrons suivant les états. La figure 5.2, *b* montre la répartition des électrons

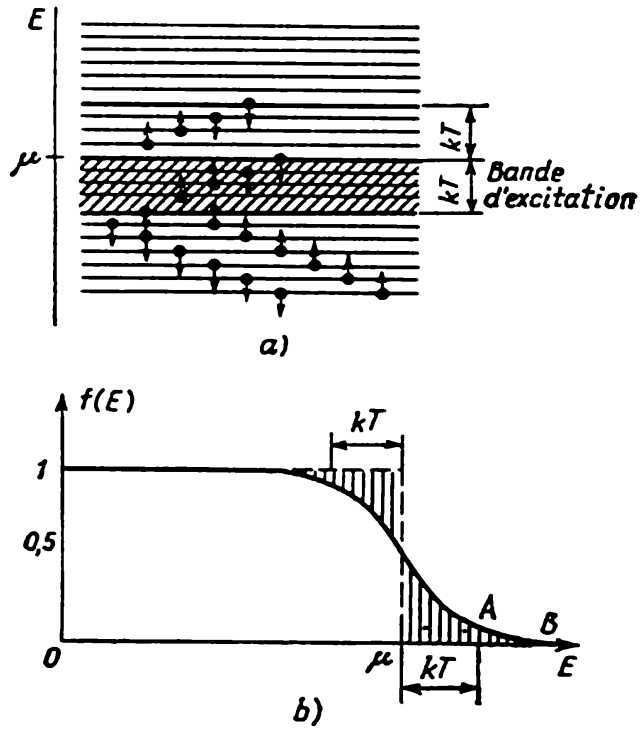

Fig. 5.2. Excitation des électrons au-dessus du niveau de Fermi μ à $T > 0$ K (*a*) et fonction de répartition de Fermi-Dirac à $T > 0$ K (*b*).

entre divers états pour $T > 0$ K. Comme il est visible sur la figure 5.2, *b*, l'élévation de la température a pour effet des répartitions étendues à une profondeur de $\sim kT$ et l'apparition d'une « queue » de la répartition qui se situe à droite de μ. Plus élevée est la température, plus forte est la variation subie par la fonction de répartition.

§ 5.2. Fonction de densité des états

Déterminons maintenant la fonction de densité des états des électrons $g(E)$. Comme les valeurs permises des vecteurs d'onde **k** des électrons sont, dans chaque bande énergétique du cristal, les mêmes que celles des vecteurs d'onde des ondes élastiques dans le réseau cristallin, la fonction de densité des états des électrons $g(E)$ est analogue à la fonction de densité de répartition des modes en fréquence $g(\omega)$ (v. § 3.3). Notamment, le nombre de différentes valeurs de **k** comprises dans l'intervalle de k à $k + dk$ est égal, suivant (3.13), à $Vk^2dk/(2\pi^2)$. Une seule différence est qu'à chaque valeur

permise de **k** correspondent deux états de l'électron ayant des spins opposés et non pas trois modes de différentes polarisations pour les ondes élastiques. De ce fait, le nombre d'états dans l'intervalle de longueurs des vecteurs d'onde dk est égal à

$$dg = Vk^2 dk/\pi^2. \tag{5.3}$$

En outre, les relations de dispersion pour l'électron (E (k)) diffèrent de façon notable de celles pour les ondes élastiques (ω (k)). Notamment pour le niveau inférieur d'une bande d'énergie (de la

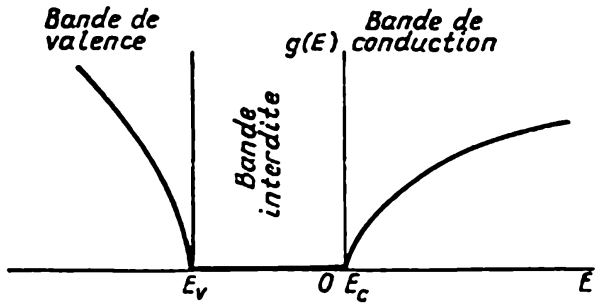

Fig. 5.3. Fonction de densité des états près du niveau inférieur de la bande de conduction et au sommet de la bande de valence.

bande de conduction, par exemple) la courbe de dispersion peut toujours être représentée sous la forme d'une parabole (fig. 4.7, a):

$$E = \hbar^2 k^2/(2m_n), \tag{5.4}$$

où m_n est la masse effective de l'électron. En tirant de (5.4)

$$k = \sqrt{\frac{2m_n E}{\hbar^2}}; \quad k\,dk = \frac{m_n}{\hbar^2}\,dE$$

et en portant ces valeurs dans (5.3), on obtient

$$dg = \frac{V}{\pi^2}\,\frac{\sqrt{2m_n^3 E}}{\hbar^3}\,dE$$

ou

$$g(E) = \frac{V}{\pi^2}\,\frac{\sqrt{2m_n^3 E}}{\hbar^3}. \tag{5.5}$$

Dans ces formules l'énergie est comptée à partir du niveau inférieur de la bande.

Près du sommet de la bande d'énergie la courbe de dispersion est aussi voisine d'une parabole. Notamment dans le voisinage du sommet de la bande de valence d'un semiconducteur on a

$$E(k) = -E_g - \frac{\hbar^2 k^2}{2m_p}, \tag{5.6}$$

où m_p est la masse effective du trou.

En se servant de cette formule, on peut obtenir l'expression suivante pour la densité des états situés près du sommet de la bande de valence :

$$g(E) = \frac{V}{\pi^2} \frac{\sqrt{2m_p^3 (E_v - E)}}{\hbar^3}, \tag{5.7}$$

où $E_v = -E_g$ est l'énergie du sommet de la bande de valence.

Les fonctions (5.5) et (5.7) sont représentées sur la figure 5.3.

§ 5.3. Concentrations d'électrons et de trous dans un semiconducteur

Concentration d'électrons dans un semiconducteur. La concentration d'électrons dans la bande de conduction peut être obtenue en intégrant la fonction de répartition complète (5.1) sur toutes les énergies de la bande de conduction :

$$n = \int\limits_{\text{(sur la bande de conduction)}} n(E)\,\mathrm{d}E = \int\limits_{0}^{E_{\text{sommet de la bande}}} g(E) f(E)\,\mathrm{d}E. \tag{5.8}$$

Ici, $n(E)$ est la fonction $N(E)$ pour $V = 1$.

L'intégration peut être étendue jusqu'à l'infini parce que la fonction à intégrer $f(E)$ décroît rapidement (exponentiellement) lorsque l'énergie croît (pour $E > \mu$) tandis que la fonction $g(E)$ croît lentement (suivant une loi de puissance). En vertu de (5.2) et (5.5) on a

$$n = \frac{1}{2\pi^2} \frac{(2m_n)^{3/2}}{\hbar^3} \int\limits_0^\infty \frac{\sqrt{E}\,\mathrm{d}E}{e^{(E-\mu)/(kT)}+1}. \tag{5.9}$$

Si le niveau de Fermi μ se situe beaucoup plus bas que le niveau inférieur de la bande de conduction $E_c = 0$, de sorte que $(E - \mu) \gg kT$, l'unité figurant au dénominateur de la fonction de Fermi-Dirac peut être négligée et les énergies dans la bande de conduction auront alors pour expression

$$f(E) = \frac{1}{e^{(E-\mu)/(kT)}+1} \approx e^{\mu/(kT)} e^{-E/(kT)}. \tag{5.10}$$

Dans ce cas le degré de remplissage de tout état dans la bande de conduction est très faible :

$$f(E) \ll 1. \tag{5.11}$$

La condition (5.11) est appelée *condition de non-dégénérescence* du gaz électronique. Elle est satisfaite si le niveau de Fermi μ se situe au moins de (2 à 3) kT plus bas que le niveau inférieur de la bande de conduction.

La fonction de répartition complète des électrons entre divers niveaux d'énergie pour le gaz non dégénéré prend alors la forme suivante :

$$n(E)\,\mathrm{d}E \approx \frac{(2m_n)^{3/2}}{2\pi^2\hbar^3}\sqrt{E}\,\mathrm{e}^{\mu/(kT)}\mathrm{e}^{-E/(kT)}\,\mathrm{d}E. \qquad (5.12)$$

C'est l'expression de la *répartition* dite *de Boltzmann*. Le nombre total de particules d'un gaz non dégénéré est proportionnel à $\mathrm{e}^{\mu/(kT)}$ et leur répartition relative suivant les énergies est déterminée par les facteurs $\sqrt{E}\mathrm{e}^{-E/(kT)}$. Elle ne dépend pas du nombre total de particules mais dépend de la température du gaz. La fonction de répartition $f(E)$ pour un gaz non dégénéré est représentée sur la figure 5.4, *a* et la fonction de répartition complète $n(E)$ sur la figure 5.4, *b*. La courbe en traits interrompus représente ici la fonction de densité des états $g(E)$. Pour un gaz non dégénéré la courbe de $n(E)$ passe toujours nettement plus bas que la courbe de $g(E)$. Elle atteint un maximum pour une énergie $E = kT/2$ et décroît pour $E > kT$. L'intégration de (5.12) sur l'énergie donne le nombre total n d'électrons en fonction de l'énergie de Fermi μ :

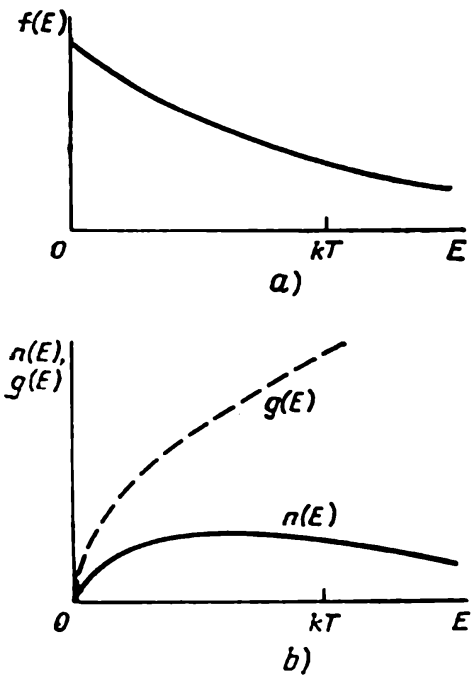

Fig. 5.4. Fonction de répartition $f(E)$ (*a*), fonction de répartition complète $n(E)$ et fonction de densité du nombre d'états $g(E)$ (*b*) pour un gaz non dégénéré.

$$n = \frac{2(2\pi m_n kT)^{3/2}}{h^3}\mathrm{e}^{\mu/(kT)} = N_c\mathrm{e}^{\mu/(kT)}. \qquad (5.13)$$

Puisqu'en vertu de (5.10), $\mathrm{e}^{\mu/kT} = f(0)$, c'est-à-dire exprime la densité moyenne de remplissage par les électrons des états confondus avec le niveau inférieur de la bande de conduction, N_c doit représenter le nombre d'états qui donnent une concentration du gaz électronique égale à n lorsqu'ils sont remplis avec la densité $\mathrm{e}^{\mu/kT}$. C'est pourquoi la quantité N_c est appelée *nombre efficace d'états ramené au niveau inférieur de la bande de conduction*. L'ordre de gran-

deur de N_c est proche du nombre d'états dans l'intervalle d'énergie $\approx kT$ à partir du niveau inférieur de la bande, c'est-à-dire du nombre d'états remplis par les électrons de façon tant soit peu substantielle. L'énergie de Fermi μ pour un gaz non dégénéré est une quantité négative. Plus bas est le niveau de Fermi, plus petit est le nombre d'électrons dans la bande de conduction à une même température.

Variation de la concentration de trous en fonction de l'énergie de Fermi. La probabilité $f_p(E)$ pour que l'état d'énergie E ne soit pas occupé par un électron a pour valeur

$$f_p(E) = 1 - f(E) = \frac{1}{1 + e^{(\mu - E)/(kT)}}. \tag{5.14}$$

Le gaz de trous dans la bande de valence sera non dégénéré, c'est-à-dire $f_p(E) \ll 1$, si la distance du niveau de Fermi μ au sommet de la bande de valence E_v est égale au moins à 2 ou 3 fois la quantité kT. Dans un tel cas

$$f_p(E) \approx e^{-(\mu - E)/(kT)} = e^{-(\mu - E_v)/(kT)} e^{-(E_v - E)/(kT)}. \tag{5.15}$$

On voit que la population en trous des états dans la bande de valence décroît exponentiellement lorsqu'on pénètre dans la profondeur de la bande de valence.

La concentration de trous s'obtient en intégrant la fonction de répartition complète des trous $f_p(E)\, g(E)$ sur l'énergie:

$$p = \frac{2(2\pi m_p kT)^{3/2}}{h^3} e^{-(\mu - E_v)/(kT)} = N_v e^{-(\mu + E_g)/(kT)}. \tag{5.16}$$

Ici N_v est le *nombre efficace d'états de la bande de valence ramené au sommet de cette bande.*

Si la condition de non-dégénérescence du gaz d'électrons dans la bande de conduction ou du gaz de trous dans la bande de valence n'est pas satisfaite, les intégrales du type (5.9) sont à calculer numériquement ou à l'aide de tables.

§ 5.4. Position du niveau de Fermi et concentration de porteurs de charges libres dans les semiconducteurs intrinsèques

La position du niveau de Fermi dans un semiconducteur intrinsèque peut être déterminée en partant de la condition que le nombre d'électrons dans la bande de conduction doit être égal au nombre de trous dans la bande de valence:

$$n_i = p_i. \tag{5.17}$$

L'indice i désigne ici et plus loin l'appartenance au semiconducteur intrinsèque.

La condition (5.17) a pour effet que le niveau de Fermi doit se situer approximativement au milieu de la bande interdite. Si le niveau de Fermi se dispose plus près du niveau inférieur de la bande de conduction, le nombre d'électrons dans un tel semiconducteur sera beaucoup plus grand que celui de trous parce que le degré de remplissage ($f(E)$) des états près du niveau inférieur de la bande de conduction est sensiblement plus élevé que le degré de leur non-remplissage ($1 - f(E)$) près du sommet de la bande de valence. La fonction de répartition $f(E)$ est représentée pour ce cas sur la figure 5.5, *a*, où l'axe des énergies est dirigé vers le haut pour rendre

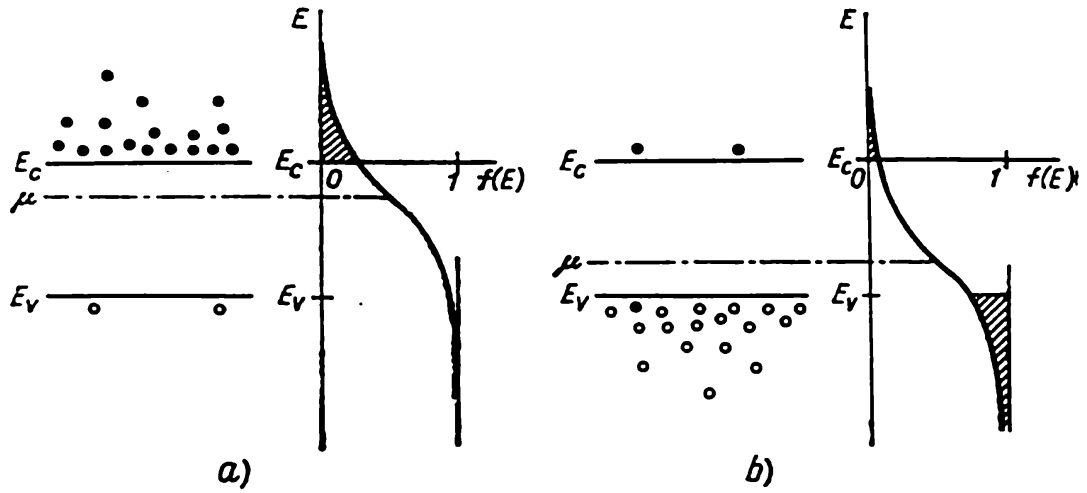

Fig. 5.5. Fonctions de répartition $f(E)$ dans des semiconducteurs de type n ($n \gg p$) (*a*) et de type p ($p \gg n$) (*b*).

plus commode la comparaison avec les diagrammes énergétiques. Au contraire, si le niveau de Fermi se situe près de la bande de valence (fig. 5.5, *b*) le nombre de trous dans la bande de valence sera beaucoup plus grand que le nombre d'électrons dans la bande de conduction. Ainsi, on ne peut assurer l'égalité du nombre d'électrons au nombre de trous que dans le cas où le niveau de Fermi est placé au milieu de la bande interdite. Toutefois, dans le cas général, le niveau de Fermi dans un semiconducteur intrinsèque ne se situe au milieu de la bande interdite que de façon approximative. Le fait est que les fonctions de densité des états dans la bande de conduction et dans la bande de valence peuvent différer l'une de l'autre. Par exemple, si $m_n > m_p$, alors, en vertu de (5.5) et (5.7), la densité des états est plus forte dans la bande de conduction que dans la bande de valence (pour une même distance des intervalles d'énergie considérés aux bords des bandes correspondantes). Dans un tel cas, pour égaliser les concentrations des électrons et des trous il convient d'abaisser un peu le niveau de Fermi en le rapprochant de la bande de valence. Par contre, si $m_p > m_n$, le niveau de Fermi doit être situé un peu plus haut que le milieu de la bande interdite.

Nous obtiendrons une valeur exacte de l'énergie de Fermi dans un semiconducteur intrinsèque en égalant entre elles $n(\mu)$ [expression (5.13)] et $p(\mu)$ [expression (5.16)]:

$$N_c e^{\mu_i/(kT)} = N_v e^{-(\mu_i+E_g)/(kT)}; \tag{5.18}$$

$$\mu_i = -\frac{E_g}{2} + \frac{kT}{2}\ln\frac{N_v}{N_c} = -\frac{E_g}{2} + \frac{3}{4}kT\ln\frac{m_p}{m_n}. \tag{5.19}$$

Lorsque $T \to 0$ K, le niveau de Fermi se situe au milieu de la bande interdite. Si la température augmente, il se déplace vers celle des bandes où la densité des états est plus faible (fig. 5.6). Etant donné que d'ordinaire m_n ne diffère pas fortement de m_p, on peut

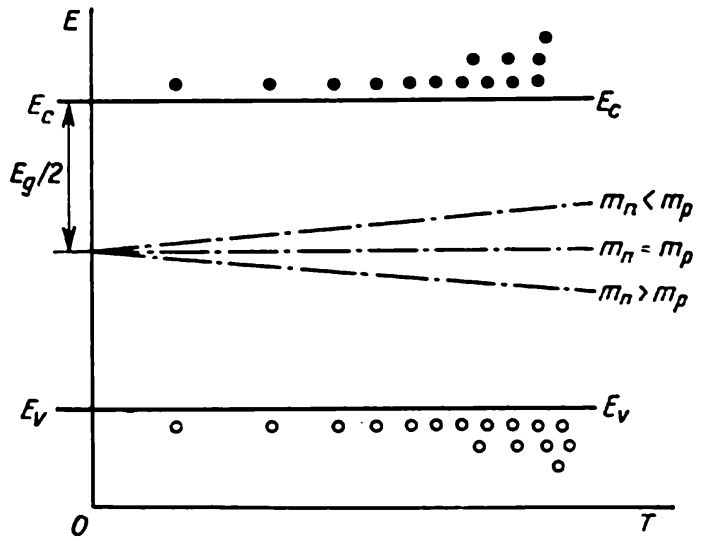

Fig. 5.6. Variation de l'énergie de Fermi (ligne en traits mixtes) en fonction de la température dans un semiconducteur intrinsèque.

considérer pour des semiconducteurs dont la largeur de la bande interdite $E_g \approx 1$ eV et plus que $\mu_i \approx -E_g/2$. Mais dans le cas des semiconducteurs à bande interdite étroite on doit tenir compte du déplacement du niveau de Fermi par rapport au milieu de la bande interdite.

En portant dans (5.13) l'expression obtenue pour μ_i, on trouve la concentration des porteurs de charges libres dans un semiconducteur intrinsèque:

$$n_i = p_i = \sqrt{N_c N_v}\, e^{-E_g/(2kT)} = \frac{2\,(2\pi\sqrt{m_n m_p}\,kT)^{3/2}}{h^3} e^{-E_g/(2kT)}. \tag{5.20}$$

L'expression (5.20) montre que la concentration d'équilibre des porteurs de courant dans un semiconducteur intrinsèque se détermine par la largeur de la bande interdite et la température du semiconducteur, la variation de n_i en fonction de T et E_g étant très brusque. C'est ainsi que la diminution de E_g de 1,12 eV (silicium) jusqu'à

0,08 eV (étain gris) a pour effet, à la température ordinaire, une augmentation de n_i de 9 ordres de grandeur; l'augmentation de la température du germanium de 100 à 600 K augmente n_i de 17 ordres de grandeur.

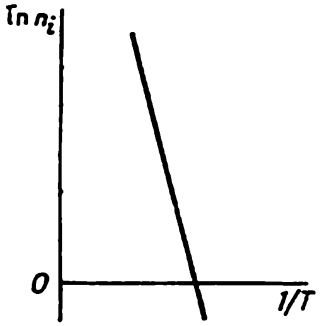

Fig. 5.7. Variation de la concentration des porteurs de courant libres en fonction de la température dans un semiconducteur intrinsèque.

Prenons le logarithme de l'expression (5.20), il vient

$$\ln n_i = \ln \sqrt{N_c N_v} - \frac{E_g}{2k} \frac{1}{T}. \tag{5.21}$$

Comme la dépendance de $\ln \sqrt{N_c N_v}$ par rapport à la température est beaucoup plus faible que celle de $1/T$, la courbe de variation de $\ln n_i$ en fonction de $1/T$ représente approximativement une droite de coefficient angulaire $-E_g/(2k)$ (fig. 5.7).

Dans les semiconducteurs à bande interdite étroite, aux températures élevées, le niveau de Fermi peut se trouver trop près (plus près que de (2 à 3) kT) de l'une ou même des deux bandes. Dans un tel cas, les expressions de n (μ) et de p (μ) pour les gaz non dégénérés d'électrons et de trous cessent d'être valables et l'équation (5.17) doit être résolue numériquement.

§ 5.5. Statistique des électrons dans des semiconducteurs extrinsèques

Dans un semiconducteur extrinsèque les électrons peuvent passer dans la bande de conduction non seulement de la bande de valence mais aussi des niveaux donneurs et les trous peuvent apparaître par suite des passages des électrons de la bande de valence aux niveaux accepteurs.

L'expression donnant la concentration des électrons au niveau d'impureté peut être obtenue dans la statistique de Fermi-Dirac si la concentration des atomes d'impureté N_{imp} est multipliée par une fonction analogue à (5.2):

$$f(E_{imp}) = \frac{1}{1 + g \exp[(E_{imp} - \mu)/(kT)]}, \tag{5.22}$$

où E_{imp} est l'énergie du niveau d'impureté et g le *facteur de dégénérescence* dépendant de la nature du centre d'impuretés. Pour des niveaux donneurs peu profonds $g = 1/2$ et pour des niveaux accepteurs peu profonds $g = 2$.

Si une impureté donatrice introduite dans le semiconducteur est à faible énergie d'ionisation E_d, le nombre d'électrons passant aux basses températures dans la bande de conduction depuis les niveaux donneurs (fig. 5.8, *a*) peut dépasser de beaucoup d'ordres de grandeur le nombre d'électrons excités depuis la bande de valence. Dans un

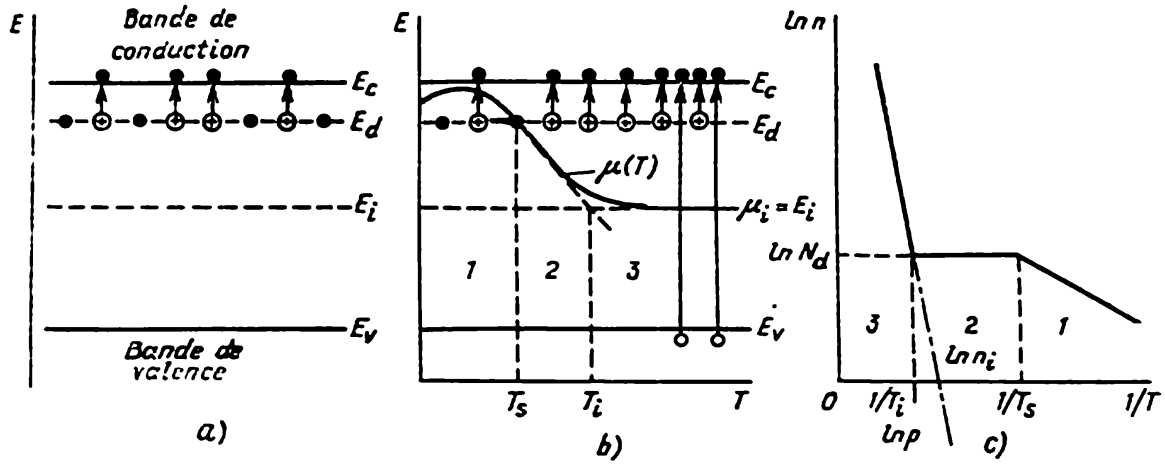

Fig. 5.8. Diagramme énergétique (*a*) et variation de l'énergie de Fermi (*b*) et de la concentration des électrons dans la bande de conduction d'un semiconducteur de type *n* (*c*) en fonction de la température (E_i — milieu de la bande interdite).

tel cas on peut négliger les transitions depuis la bande de valence et déterminer l'énergie de Fermi en égalant le nombre n d'électrons de conduction au nombre de places vacantes N_d^+ aux niveaux donneurs.

Si la concentration des donneurs n'est pas trop forte, le gaz électronique dans la bande de conduction reste non dégénéré malgré l'augmentation de la concentration des électrons de conduction par rapport à un semiconducteur intrinsèque pur. La concentration des électrons se détermine alors par l'expression (5.13).

Le nombre de places vacantes aux niveaux donneurs (l'énergie d'un niveau donneur étant égale à $-E_d$)

$$N_d^+ = N_d\,[1 - f(-E_d)] = \frac{N_d}{1 + 2\exp[(E_d + \mu)/(kT)]}. \qquad (5.23)$$

En égalant n et N_d^+, on obtient une équation pour la détermination de μ:

$$N_c \exp\frac{\mu}{kT} = \frac{N_d}{1 + 2\exp[\mu/(kT)]\exp[E_d/(kT)]}. \qquad (5.24)$$

Très basses températures. L'équation (5.24) peut être simplifiée si l'on tient compte du fait qu'aux températures suffisamment basses les donneurs sont ionisés faiblement ($N_d^+ \ll N_d$) et le dénominateur au second membre de l'expression (5.23) est beaucoup plus grand que

l'unité. Alors

$$N_d^+ = \frac{N_d}{2} \exp\left(-\frac{E_d + \mu}{kT}\right), \tag{5.25}$$

et l'équation (5.24) prend la forme

$$2N_c \exp\frac{\mu}{kT} = N_d \exp\left(-\frac{E_d}{kT}\right) \exp\left(-\frac{\mu}{kT}\right). \tag{5.26}$$

D'où

$$\mu = -\frac{E_d}{2} + \frac{kT}{2} \ln\frac{N_d}{2N_c} = -\frac{E_d}{2} + \frac{kT}{2} \ln\frac{N_d h^3}{4(2\pi m_n kT)^{3/2}}. \tag{5.27}$$

Ainsi, lorsque $T \to 0$ K, le niveau de Fermi tend vers le milieu entre le niveau inférieur de la bande de conduction et le niveau donneur (fig. 5.8, *b*). Si la température s'élève à partir de 0 K, l'énergie de Fermi commence par augmenter parce qu'aux températures suffisamment basses le dénominateur de la fraction sous le signe logarithme dans (5.27) est inférieur au numérateur. Lorsque la température augmente encore, l'énergie de Fermi atteint son maximum, commence à décroître et, à une certaine température T_s, le niveau de Fermi traverse le niveau des donneurs $-E_d$. Au voisinage de cette température l'expression (5.27) devient inapplicable parce qu'elle a été obtenue dans l'hypothèse que $N_d^+ \ll N_d$, alors que si le niveau de Fermi se confond avec le niveau des donneurs, $N_d^+ = N_d/3$. Pour la détermination de μ dans ce domaine de température il convient d'avoir recours à l'équation plus générale (5.24).

La température T_s est appelée *température de déplétion de l'impureté*. Elle peut être déterminée à partir de l'équation

$$N_c \exp\frac{\mu}{kT_s} = N_c \exp\left(-\frac{E_d}{kT_s}\right) = \frac{N_d}{3}. \tag{5.28}$$

D'où

$$T_s = \frac{E_d}{k \ln(3N_c/N_d)}. \tag{5.29}$$

Cette expression montre que la température de déplétion est d'autant plus élevée que l'énergie d'ionisation E_d de l'impureté est plus grande ce qui est parfaitement naturel. Ce qui apparaît surprenant, de prime abord, c'est la variation de la température de déplétion en fonction de la concentration de l'impureté. Cela s'explique par le fait qu'à toute température il se produit un *échange* d'électrons entre le niveau donneur et les niveaux inférieurs de la bande de conduction situés dans l'intervalle d'énergie de l'ordre de quelques kT. Plus grand est le nombre de donneurs, plus forte est la probabilité qu'ils captent un électron de la bande de conduction et donc plus élevée est la température de déplétion T_s de l'impureté. Les températures de déplétion sont peu élevées. Par exemple, pour une impureté de $E_d = 0{,}01$ eV dans le germanium, $T_s = 30$ K pour $N_d = 10^{16}$ cm^{-3}.

La concentration des électrons dans la bande de conduction aux très basses températures s'obtient en portant dans (5.13) l'expression (5.27) de μ:

$$n = \sqrt{\frac{N_c N_d}{2}} \exp\left(-\frac{E_d}{2kT}\right). \tag{5.30}$$

En prenant le logarithme de cette expression, on trouve

$$\ln n = \ln \sqrt{\frac{N_c N_d}{2}} - \frac{E_d}{2k} \frac{1}{T}. \tag{5.31}$$

La dépendance logarithmique par rapport à la température du premier terme du second membre de cette expression étant beaucoup plus faible que $1/T$, la variation de $\ln n$ en fonction de $1/T$ dans le domaine de très basses températures est approximativement linéaire avec le coefficient angulaire $-E_d/(2k)$ (région *1* de la fig. 5.8, *c*).

Domaine de déplétion de l'impureté (région *2* des figures 5.8, *b*, *c*). Lorsque la température s'élève encore au-delà de T_s, pratiquement tous les électrons passent des atomes d'impureté dans la bande de conduction et alors

$$n = N_d.$$

L'énergie de Fermi dans ce domaine des températures s'obtient à partir de l'équation

$$N_d = N_c \exp[\mu/(kT)]. \tag{5.32}$$

Elle a pour valeur

$$\mu = kT \ln \frac{N_d}{N_c} = kT \ln \frac{N_d h^3}{2(2\pi m_n kT)^{3/2}}. \tag{5.33}$$

Comme N_c devient plus grand que N_d même pour $T < T_s$, il résulte de (5.33) que le niveau de Fermi baisse lorsque la température augmente.

Domaine de la conductibilité intrinsèque. Si la température continue à augmenter, ce sont les électrons passant de la bande de valence dans la bande de conduction qui commencent à apporter une contribution notable à la concentration. La concentration des électrons croît de nouveau (région *3* de la figure 5.8, *c*) et finit par devenir pratiquement égale à la concentration n_i dans un semiconducteur intrinsèque. Pour la température T_i de passage à la conductibilité intrinsèque, on prend la température à laquelle la concentration de porteurs de charges dans le semiconducteur intrinsèque est égale à la concentration des électrons dans un semiconducteur du type n dans le domaine de déplétion de l'impureté:

$$n_i(T_i) = N_d,$$

c'est-à-dire, en vertu de (5.20),

$$\sqrt{N_c N_v} \exp\left(-\frac{E_g}{2kT_i}\right) = N_d.$$

D'où

$$T_i = \frac{E_g}{k \ln (N_c N_v / N_d^2)}. \tag{5.34}$$

Cette équation montre que la température de passage à la conductibilité intrinsèque est d'autant plus élevée que la bande interdite du semiconducteur E_g est plus large et la concentration d'impureté N_d est plus forte. Par exemple, pour $N_d = 10^{16}$ cm^{-3} la température

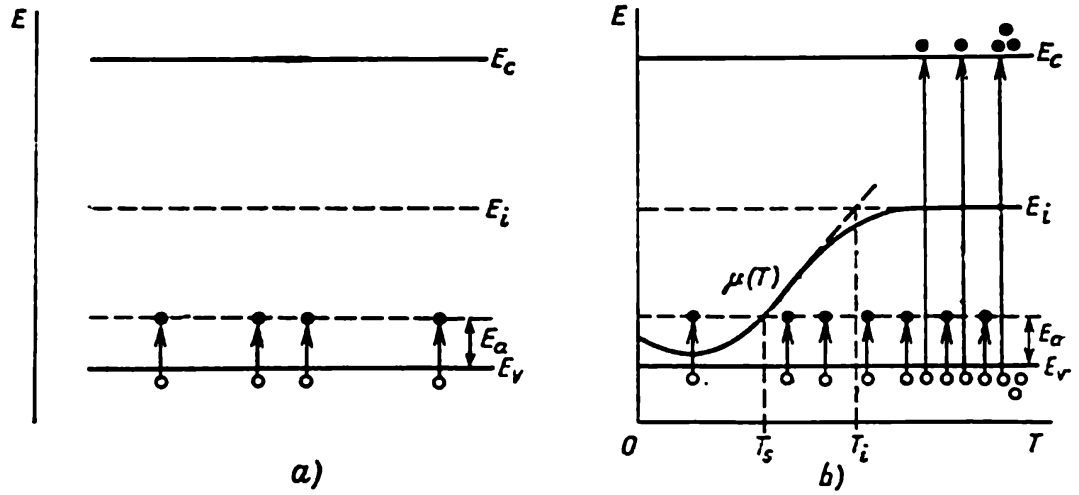

Fig. 5.9. Diagramme énergétique (*a*) et variation de l'énergie de Fermi en fonction de la température (*b*) pour un semiconducteur de type *p* (E_i — milieu de la bande interdite).

T_i est égale à 480 K pour le germanium et à près de 650 K pour le silicium. La valeur de T_i détermine la température limite de fonctionnement de la plupart des appareils à semiconducteurs.

Semiconducteur du type *p* (fig. 5.9, *a*). Les formules donnant l'énergie de Fermi et la concentration des porteurs de charges dans un semiconducteur dopé aux accepteurs sont obtenues de la même façon que pour le semiconducteur du type *n*. Elles ont la même forme que celle pour un semiconducteur du type *n* à condition qu'au lieu de N_c on y écrit N_v, au lieu de E_d, E_a et au lieu de l'énergie de Fermi μ comptée à partir du niveau inférieur de la bande de conduction vers le haut sur l'échelle des énergies, on écrit μ' — l'énergie de Fermi comptée vers le bas à partir du sommet de la bande de valence. La courbe de variation de l'énergie de Fermi en fonction de la température dans un semiconducteur du type *p* est montrée sur la figure 5.9, *b*. Quant à la courbe représentative de ln *p* en fonction de $1/T$, elle est tout à fait analogue à la courbe correspondante pour le semiconducteur du type *n* (v. fig. 5.8, *c*).

Loi d'action de masse. En multipliant la concentration des électrons (5.13) par celle des trous (5.16), on obtient une relation importante

$$np = N_c N_v e^{-E_g/(kT)}. \tag{5.35}$$

Le rapprochement entre (5.35) et (5.20) donne

$$np = n_i^2 (T). \tag{5.36}$$

Les formules obtenues expriment la *loi d'action de masse.* Elle peut être énoncée comme suit: *le produit des concentra' ns des électrons et des trous dans un semiconducteur ne dépend pas d on dopage, il dépend uniquement de la température; ce produit est é l au carré de la concentration de porteurs de charges dans un semico ucteur intrinsèque.*

Prenons le logarithme de (5.36), il vient

$$(\ln n + \ln p)/2 = \ln n_i. \tag{5.37}$$

Cette relation a été utilisée pour la construction des graphiques de la figure 5.8, *c*, où la ligne en traits mixtes représente $\ln n_i$ et la ligne en traits interrompus $\ln p$. Dans le domaine de la conductibilité d'impureté la concentration de porteurs minoritaires (de trous dans un semiconducteur du type *n*) est beaucoup plus faible que dans un semiconducteur intrinsèque. Cela provient du fait que les concentrations d'équilibre s'établissent non seulement par suite de génération de porteurs mais également par suite de leur recombinaison. Dans le cas où, en plus des électrons ayant passé de la bande de valence, la bande de conduction contient un grand nombre d'électrons d'impureté, les niveaux libérés dans la bande de valence sont réoccupés beaucoup plus vite par les électrons, ce qui a pour effet une brusque diminution de la concentration de trous.

Il convient de souligner que les relations (5.13) et (5.16) n'étant valables que pour des gaz non dégénérés, la loi d'action de masse elle aussi n'est vraie que pour des semiconducteurs non dégénérés.

Semiconducteurs fortement dopés. Lorsque le degré de dopage augmente, la courbe de variation de $\ln n$ en fonction de $1/T$ dans le domaine de la conductibilité d'impureté se déplace vers le haut. Le coefficient angulaire de la droite dans le domaine de très basses températures reste inchangé parce qu'il se détermine par l'énergie d'ionisation E_d d'un atome donneur distinct et ne dépend pas de leur concentration. Toutefois, il n'en est pas ainsi jusqu'à l'instant où la concentration d'impuretés atteint un certain niveau au-delà duquel l'inclinaison de la droite $\ln n$ sur l'axe des $1/T$ commence à diminuer (courbe de N_d'' de la figure 5.10, *a*) pour s'annuler finalement (courbe de N_d''').

La diminution de l'inclinaison des droites considérées commence pour des concentrations d'impureté lorsque les fonctions d'onde des

électrons des atomes d'impureté voisins commencent à se recouvrir. Alors, au lieu des niveaux d'énergie distincts pour chaque atome d'impureté (fig. 5.10, *b*) il apparaît une *bande d'énergie d'impureté* (fig. 5.10, *c*). Pour une concentration encore plus forte de l'impureté, cette bande se réunit à la bande de conduction de façon à former une bande unique d'énergies permises (fig. 5.10, *d*). L'apparition de la bande d'impureté a pour effet de réduire l'énergie d'activation de la conductibilité d'impureté E_d et, lorsque cette bande se réunit à la bande de conduction, l'énergie d'activation tombe à zéro.

La variation de la densité des états en fonction de l'énergie près du niveau inférieur de la bande de conduction après la réunion avec

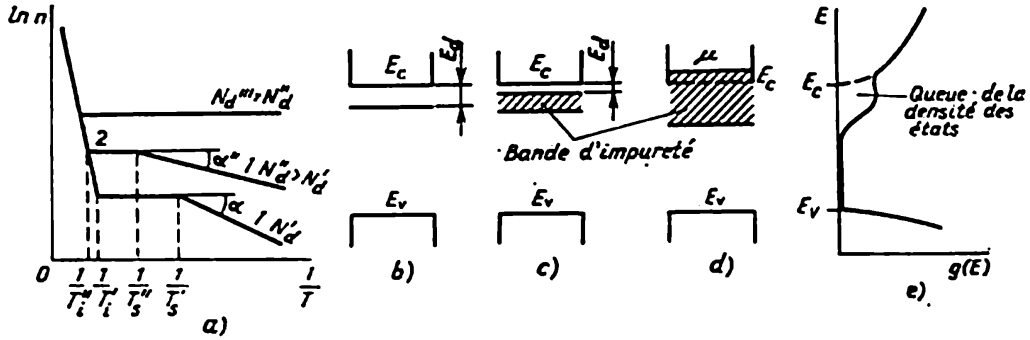

Fig. 5.10. Variation de la concentration des électrons dans la bande de conduction en fonction de la température pour des semiconducteurs contenant des quantités différentes de l'impureté donatrice (*a*); niveau d'impureté dans la bande interdite (*b*); formation de la bande d'impureté à partir d'un niveau d'impureté en cas de forte concentration de l'impureté (*c*); recouvrement de la bande d'impureté et de la bande de conduction (*d*) et fonction de densité des états (*e*) dans des semiconducteurs dégénérés.

la bande d'impureté prend la forme montrée sur la figure 5.10, *e*. Il se forme une *queue de la densité des états*.

La formation de la bande d'impureté se produit de la même façon que celle des bandes d'énergie à partir des niveaux étroits lors du rapprochement des atomes de la substance principale (v. chap. 4). Toutefois le niveau d'impureté s'étale en une bande de largeur notable lorsque la distance entre les atomes d'impureté voisins est encore beaucoup plus grande que les distances interatomiques pour les atomes de la substance de base. Cela tient à ce que l'« orbite » d'un électron faiblement lié à l'atome d'impureté a un très grand rayon et le recouvrement des couches électroniques des atomes d'impureté se produit pour une distance suffisamment grande qui les sépare. Par exemple, dans le germanium l'énergie d'ionisation des donneurs peu profonds diminue de façon appréciable à partir de la concentration $N_d \approx 10^{16}$ cm^{-3}, lorsque les distances entre les atomes d'impureté voisins sont approximativement égales à 50 nm.

Le gaz électronique dans un semiconducteur fortement dopé peut être dégénéré surtout à des températures très inférieures à la tempé-

rature de passage à la conductibilité intrinsèque T_i. Au fur et à mesure que la température du semiconducteur s'approche de T_i, le niveau de Fermi tend vers le milieu de la bande interdite et, si le semiconducteur n'est pas à bande interdite étroite, la dégénérescence est levée.

Les niveaux accepteurs s'étalent eux aussi en une bande lorsque leur concentration est suffisamment forte.

Semiconducteurs compensés. Bien souvent un semiconducteur contient à la fois une impureté donatrice et une impureté acceptrice

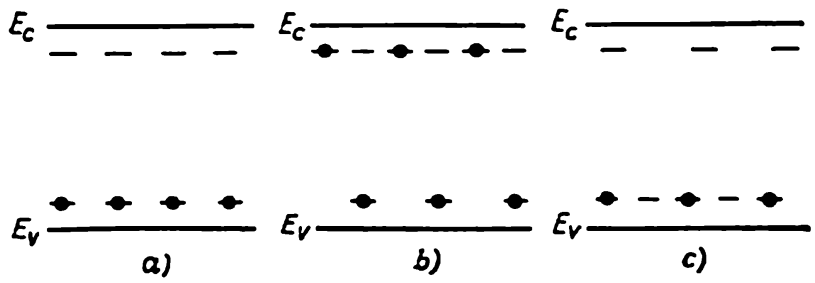

Fig. 5.11. Diagrammes énergétiques des semiconducteurs compensés: a — $N_d = N_a$; b — $N_d > N_a$; c — $n_d < N_a$.

en des quantités comparables. De tels semiconducteurs sont dits *compensés.*

Commençons par examiner le cas d'une compensation complète lorsque la concentration N_d de donneurs est égale à celle N_a d'accepteurs. Le diagramme énergétique d'un tel semiconducteur est montré sur la figure 5.11, *a*. Suivant la statistique de Fermi-Dirac, lorsque $T \to 0$ K, les électrons occupent les niveaux d'énergie les plus bas. Dans un tel cas, tous les niveaux accepteurs seront occupés par les électrons et tous les niveaux donneurs seront libres. Lorsque la température s'élève, les électrons passent de la bande de valence dans la bande de conduction et pratiquement c'est seulement grâce à de tels passages qu'apparaissent des porteurs de charges libres — les électrons et les trous. Cela signifie que dans un tel semiconducteur les concentrations de porteurs libres sont les mêmes que dans un semiconducteur intrinsèque.

Si $N_d > N_a$ (fig. 5.11, *b*) et $T \to 0$ K, tous les états accepteurs sont occupés par les électrons et si la température s'élève ils ne pourront pas, de même que dans le cas considéré plus haut, recevoir les électrons de la bande de valence. Mais à la différence de la compensation complète, $(N_d - N_a)$ donneurs conservent à $T = 0$ K leurs électrons et peuvent les fournir à la bande de conduction en cas d'élévation de la température. Un tel semiconducteur se comporte donc comme un semiconducteur du type n avec une concentration de donneurs égale à $(N_d - N_a)$.

Si $N_d < N_a$ (fig. 5.11, *c*), le semiconducteur considéré se comporte comme un semiconducteur du type *p* avec une concentration d'accepteurs égale à $(N_a - N_d)$.

En réalité, aux basses températures, il existe certaines différences dans la position du niveau de Fermi et les concentrations de porteurs de charges par rapport aux semiconducteurs non compensés. Mais nous n'allons pas les examiner ici.

§ 5.6. Statistique des électrons dans les métaux

On sait (v. § 4.5) que dans les métaux la bande la plus haute parmi les bandes d'énergie occupées par les électrons, à savoir la bande de conduction, n'est remplie qu'en partie.

A $T = 0$ K, les électrons occupent tous les états inférieurs de la bande de conduction (un électron dans chaque état) jusqu'au niveau de Fermi μ (fig. 4.8, *a*, *b*). Comme le nombre d'électrons de conduction dans un métal est connu (pour les métaux monovalents par exemple il est égal au nombre d'atomes), l'énergie μ peut être déterminée en comptant à partir du niveau inférieur de la bande le nombre d'états égal au nombre d'électrons :

$$n = \int_0^{\mu} g(E)\,\mathrm{d}E. \tag{5.38}$$

Ici, n est la concentration d'électrons.

Dans les métaux, la fonction de densité des états $g(E)$ a la forme (5.5) non seulement près du niveau inférieur mais presque pour toute la bande de conduction et m_n coïncide pratiquement avec la masse m_e d'un électron libre. En portant (5.5) dans (5.38), on obtient

$$n = \frac{1}{3\pi^2}\left(\frac{2m_e\mu}{\hbar^2}\right)^{3/2}. \tag{5.39}$$

D'où

$$\mu = \frac{\hbar^2}{2m_e}(3\pi^2 n)^{2/3}. \tag{5.40}$$

Dans le Li par exemple, le calcul à l'aide de la formule (5.40) donne $\mu = 4{,}72$ eV alors que la valeur expérimentale est $\mu = 5{,}22$ eV, de sorte que l'accord n'est pas mauvais si l'on tient compte des hypothèses simplificatrices faites plus haut.

Connaissant l'énergie de Fermi μ, on peut calculer la vitesse maximale de mouvement des électrons à l'intérieur du cristal :

$$v_{max} = \sqrt{2\mu/m_e}. \tag{5.41}$$

Elle vaut $\sim 10^6$ m/s. Et cela à 0 K lorsque, suivant la théorie classique, tout mouvement doit cesser ! C'est le résultat de l'action

des lois du mouvement purement quantiques — du *principe de Pauli* qui stipule que les électrons possédant des vitesses si grandes ne peuvent pas les perdre sinon ils devraient descendre sur des niveaux d'énergie moindre déjà occupés. Comme il sera montré plus loin, un tel comportement non classique du gaz électronique dans le métal influe sur la conductibilité électrique, la conductibilité thermique, la chaleur massique et autres caractéristiques des métaux ainsi que sur leur variation en fonction de la température. L'anomalie dans

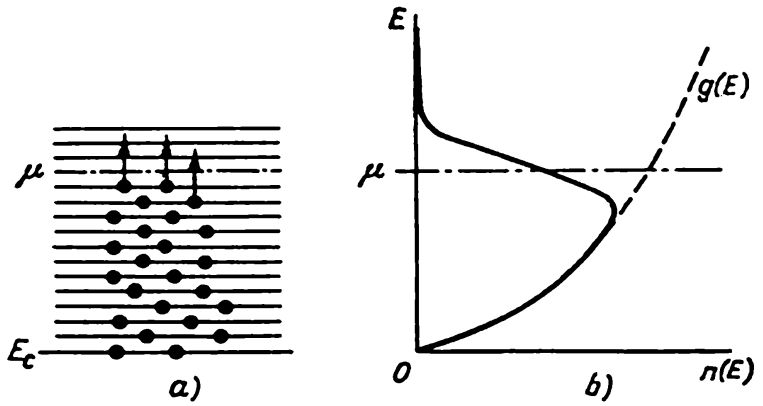

Fig. 5.12. Excitation thermique des électrons dans le métal (*a*) et leur répartition $n(E)$ suivant les énergies à $T > 0$ K (*b*).

le comportement du gaz électronique est observée non seulement dans les métaux mais aussi dans des semiconducteurs fortement dopés et, en général, dans toute collectivité d'électrons dans laquelle sont complètement occupés au moins les états énergétiques inférieurs, c'est-à-dire pour lesquels

$$f(E_c) \approx 1.$$

Un tel gaz électronique est dit *dégénéré*.

A des températures supérieures à 0 K les électrons occupant les états dans l'intervalle d'énergie $\sim kT$ au voisinage du niveau de Fermi passent à des niveaux plus hauts (fig. 5.12, *a*). La répartition des électrons entre divers niveaux d'énergie est illustrée par la figure 5.12, *b*. Comme kT est toujours très inférieur à μ (à la température ordinaire par exemple, $kT \approx 0{,}025$ eV), l'énergie varie, lorsque la température s'élève, seulement pour une petite partie des électrons.

La position du niveau de Fermi à $T > 0$ K peut être déterminée à partir de la condition que le nombre total des électrons de conduction doit rester inchangé quelle que soit la température:

$$n = \int_0^\infty n(E)\,\mathrm{d}E = \int_0^\infty g(E)\,f_\mu(E)\,\mathrm{d}E. \qquad (5.42)$$

Dans la dernière intégrale, l'indice μ montre que la fonction de répartition $f(E)$ dépend de l'énergie de Fermi comme d'un paramètre. Cette intégrale ne peut pas être représentée sous la forme d'une fonction analytique. Une valeur approchée de l'énergie de Fermi μ est égale à

$$\mu(T) = \mu(0)\left[1 - \frac{\pi^2}{12}\left(\frac{kT}{\mu(0)}\right)^2\right], \tag{5.43}$$

où $\mu(0)$ est l'énergie de Fermi à $T = 0$ K calculée au moyen de la formule (5.40).

L'expression (5.43) montre que l'énergie de Fermi varie très peu avec la température. C'est ainsi par exemple à la température ordinaire pour un métal de $\mu(0) = 5$ eV la variation de μ est de $\sim 10^{-5}$ eV. L'énergie de Fermi diminue un peu avec la température parce que la densité des états $g(E)$ augmente au voisinage du niveau $E = \mu$. Si le niveau de Fermi ne changeait pas sa position, l'augmentation de la population des niveaux de $E > \mu$ dans l'intervalle $\Delta E \approx kT$ ne serait pas compensée par une diminution de la population des niveaux de $E < \mu$ moins serrés et le nombre total des électrons augmenterait.

Les expressions obtenues plus haut sont valables pour des métaux monovalents dans lesquels les électrons de valence n'occupent de façon incomplète qu'une bande d'énergie. La situation qui se présente dans des métaux polyvalents est plus complexe (v. § 4.8), mais ici encore le gaz électronique est dégénéré et son état varie très peu avec la température.

La faible variation de la répartition énergétique des électrons du métal en fonction de la température a pour effet que les valeurs moyennes des grandeurs caractérisant le gaz d'électrons (énergie moyenne, vitesses moyennes, etc.) ne dépendent pratiquement pas de la température. Il en résulte que le gaz électronique dans les métaux possède des caractéristiques qui sont notablement différentes des caractéristiques classiques des gaz non dégénérés. Considérons à titre d'exemple l'influence de la dégénérescence sur la chaleur massique du gaz électronique.

§ 5.7. Chaleur massique du gaz électronique dans les métaux et les semiconducteurs

En plus des ions qui constituent le réseau cristallin et vibrent autour des positions d'équilibre, les métaux contiennent aussi des électrons libres dont le nombre dans un métal monovalent est égal au nombre d'atomes. Aussi la chaleur massique C_v d'un métal doit-elle être égale à la somme de la chaleur massique du réseau $C_{rés}$ et de la chaleur massique du gaz électronique C_{el}. Evaluons l'ordre de grandeur de C_{el}.

Comme il a été indiqué au § 5.2, lorsque le métal est échauffé à une température T, c'est seulement une petite partie d'électrons dans l'intervalle d'énergie $\sim kT$ au voisinage du niveau de Fermi μ qui subit une excitation thermique. Si l'on admet, pour simplifier le problème, que la répartition des niveaux d'énergie dans la bande de conduction est uniforme, la part des électrons excités sera égale approximativement à kT/μ. En considérant que chacun de ces électrons a reçu une énergie d'excitation égale en moyenne à $^3/_2\, kT$, on obtient la variation de l'énergie du gaz électronique E_{el} produite par son échauffement à la température T :

$$E_{el} = {}^3/_2\, N\, (kT)^2/\mu, \tag{5.44}$$

où N est le nombre d'électrons de conduction du métal. La chaleur massique d'un tel gaz a pour valeur

$$C_{el} \approx 3Nk\, (kT/\mu). \tag{5.45}$$

Une théorie plus rigoureuse donne pour C_{el} l'expression

$$C_{el} = \frac{\pi^2}{2} Nk \left(\frac{kT}{\mu}\right). \tag{5.46}$$

Notons que suivant la théorie classique chaque électron libre dans le métal doit avoir une énergie cinétique moyenne de $^3/_2\, kT$ et que l'énergie totale du gaz électronique doit être égale à $^3/_2\, NkT$ et la chaleur massique $C_{el} = {}^3/_2\, Nk$, c'est-à-dire 50 % de la chaleur massique du réseau cristallin [v. (3.27)]. En comparant (5.46) avec (3.27) et en tenant compte que pour les métaux l'énergie μ se mesure par plusieurs électrons-volts tandis que $kT = 0{,}025$ eV à $T = 300$ K, on obtient $C_{el} \approx 0{,}01 C_{rés}$.

Ainsi, du fait qu'aux températures ordinaires seule une partie peu importante d'électrons libres du métal subit une excitation thermique, la chaleur massique du gaz électronique ne s'élève qu'à quelques pour cent de la chaleur massique du réseau, ce qui est une conséquence de la dégénérescence du gaz électronique dans les métaux. Dans les semiconducteurs (pour fixer les idées, nous raisonnerons sur les semiconducteurs du type n) la chaleur massique du gaz électronique dépend dans le cas général non seulement de l'énergie moyenne des électrons dans la bande de conduction mais aussi de leur concentration. Dans le domaine de température de déplétion de l'impureté la concentration des électrons dans la bande de conduction est indépendante de la température et est égale à la concentration d'impureté N_d. Si, dans ce cas, le semiconducteur n'est pas dopé fortement de sorte que le gaz électronique n'est pas dégénéré, sa chaleur massique est donnée par l'expression classique

$$C_{el} = {}^3/_2\, N_d k. \tag{5.47}$$

Pour un gaz dégénéré la chaleur massique dans le domaine considéré est moindre que celle donnée par (5.47). Dans le cas d'une forte dégénérescence, lorsque $\mu \gg kT$, la chaleur massique électronique est bien décrite par la formule (5.46). Il est vrai que les semiconducteurs, même fortement dopés, sont, généralement parlant, des substances suffisamment pures. C'est ainsi que même pour $N_d = 10^{20}$ cm^{-3} la concentration de l'impureté n'est égale qu'à $\sim$0,2 %. Aussi, le nombre d'électrons dans la bande de conduction est-il toujours petit par rapport au nombre d'atomes du réseau cristallin, de sorte que $C_{el} \ll C_{rés}$.

Dans le domaine de la conductibilité intrinsèque la chaleur massique électronique du semiconducteur prend une valeur nettement supérieure à celle donnée par (5.47). Cela tient à ce que dans ce domaine la concentration des porteurs de courant libres (maintenant des électrons et des trous) peut dépasser sensiblement la concentration d'atomes d'impureté ($n_i \gg N_d$) et la naissance de chaque paire nouvelle de porteurs exige une énergie $E_g \gg kT$. Dans ce domaine, la chaleur massique électronique du semiconducteur constitue une partie substantielle de sa chaleur massique totale.

CHAPITRE 6

PHÉNOMÈNES CINÉTIQUES

En partant des notions, développées plus haut, de la théorie des bandes, de la statistique des porteurs de charges libres dans les semiconducteurs et les métaux et de la théorie des vibrations thermiques du réseau cristallin, nous exposons dans ce chapitre les idées principales actuelles sur la nature de la conductibilité électrique des solides, y compris de la supraconductivité, ainsi que sur la nature des principaux phénomènes galvanomagnétiques.

§ 6.1. Répartition des électrons suivant leurs vitesses

De la fonction de répartition des électrons suivant l'énergie $n(E)\,\mathrm{d}E$ il est facile de passer à la fonction de répartition suivant le module de la vitesse en utilisant les relations

$$E = m_n v^2/2; \quad \mathrm{d}E = m_n v\,\mathrm{d}v.$$

Pour un gaz électronique *non dégénéré* on obtient à partir de (5.12)

$$n(v)\,\mathrm{d}v = \frac{8\pi m_n^3}{h^3}\,\mathrm{e}^{\mu/(kT)} v^2 \mathrm{e}^{-m_n v^2/(2kT)}\,\mathrm{d}v. \tag{6.1}$$

En exprimant $\mathrm{e}^{\mu/(kT)}$ moyennant le nombre total n des électrons à l'aide de (5.13), on obtient la *fonction de Maxwell classique pour la répartition des particules de gaz suivant leurs vitesses*:

$$n(v)\,\mathrm{d}v = 4\pi n \left(\frac{m_n}{2\pi kT}\right)^{3/2} v^2 \mathrm{e}^{-m_n v^2/(2kT)}\,\mathrm{d}v. \tag{6.2}$$

La courbe représentative de la fonction $n(v)$ est montrée sur la figure 6.1, *a*.

Pour obtenir la *répartition des électrons suivant les composantes du vecteur vitesse* il est commode de considérer l'occupation des états permis dans l'« espace des vitesses ». En utilisant la relation entre la vitesse $\mathbf{v}$ de l'électron et son vecteur d'onde $\mathbf{k}$

$$m_n \mathbf{v} = \hbar \mathbf{k},$$

on obtient par exemple pour la composante x de la vitesse

$$\delta v_x = \frac{\hbar}{m_n}\,\delta k_x$$

et alors les valeurs permises voisines des projections du vecteur **v** sur les axes v_x, v_y, v_z sont séparées par des intervalles $h/(Lm_n)$ (v. fig. 3.5). Pour un cristal de dimensions unitaires ($L = 1$) ces intervalles sont égaux à h/m_n. Ainsi le nombre total d'états dans un certain intervalle, isolé par la pensée, de vitesses $dv_x\, dv_y\, dv_z$ est égal à $2dv_x\, dv_y\, dv_z/(h^3/m_n^3)$, où le facteur 2 tient compte du fait qu'à chaque valeur permise de **v** correspondent deux états de spins opposés.

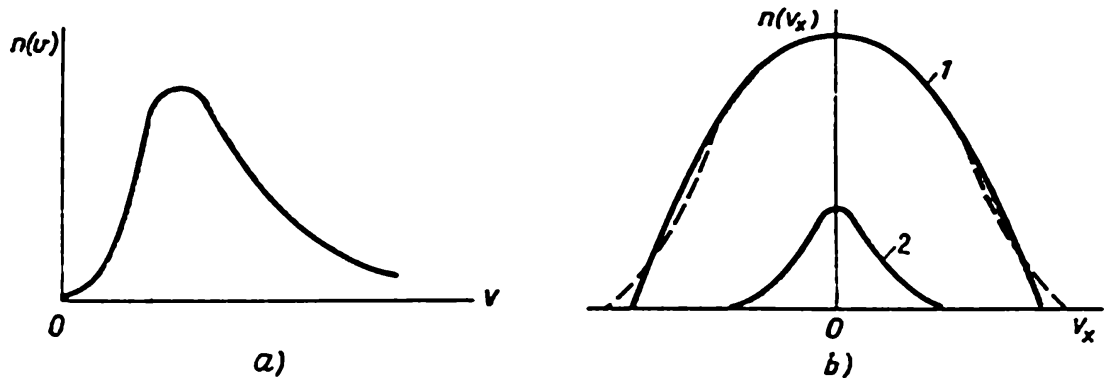

Fig. 6.1. Répartition de Maxwell des électrons suivant les vitesses (*a*); répartition des électrons suivant une des composantes (v_x) de la vitesse pour un gaz électronique dégénéré (courbe *1*) et non dégénéré (courbe *2*) (*b*).

La probabilité d'occupation de tout état dépend uniquement de son énergie. Pour un gaz non dégénéré elle a, suivant (5.10), pour valeur

$$f(E) = \exp\left(\frac{\mu}{kT}\right)\cdot\exp\left(-\frac{E}{kT}\right) =$$

$$= \exp\left(\frac{\mu}{kT}\right)\exp\left(\frac{m_n(v_x^2+v_y^2+v_z^2)}{2kT}\right). \qquad (6.3)$$

En multipliant cette fonction par le nombre d'états dans l'élément d'espace des vitesses isolé, on obtient

$$n(v_x,\ v_y,\ v_z)\, dv_x\, dv_y\, dv_z =$$

$$= \frac{2m_n^3}{h^3}\exp\left(\frac{\mu}{kT}\right)\cdot\exp\left[-\frac{m_n}{2kT}(v_x^2+v_y^2+v_z^2)\right] dv_x\, dv_y\, dv_z, \qquad (6.4)$$

où $n(v_x,\ v_y,\ v_z)$ est la fonction de densité de la répartition des électrons suivant les vitesses.

En exprimant $\exp[\mu/(kT)]$ en fonction de la concentration n d'électrons [v. formule (5.13)], on obtient

$$n(v_x,\ v_y,\ v_z) = n\left(\frac{m_n}{2\pi kT}\right)^{3/2}\exp\left[-\frac{m_n}{2kT}(v_x^2+v_y^2+v_z^2)\right]. \qquad (6.5)$$

Enfin, l'intégration de cette fonction sur v_y et v_z dans les limites de $-\infty$ à $+\infty$ donne la *fonction de répartition des électrons d'un gaz non dégénéré suivant la composante de la vitesse* v_x:

$$n(v_x)\,dv_x = dv_x \iint_{-\infty}^{+\infty} n(v_x, v_y, v_z)\,dv_y\,dv_z =$$

$$= n\left(\frac{m_n}{2\pi kT}\right)^{1/2} \exp\left(-\frac{m_n v_x^2}{2kT}\right) dv_x. \qquad (6.6)$$

Cette fonction est représentée par la courbe *2* de la figure 6.1, *b*.

A $T = 0$ K, les électrons d'un gaz dégénéré remplissent complètement tous les états à l'intérieur d'une sphère de rayon $v_F = \sqrt{2\mu/m_n}$ (fig. 6.2). Ainsi la fonction de densité de la répartition des électrons suivant les vitesses à l'intérieur de cette sphère de Fermi coïncide avec la fonction de densité des états:

$$n(v_x, v_y, v_z) = 2m_n^3/h^3. \qquad (6.7)$$

La répartition des électrons suivant l'une des composantes de la vitesse (v_x) est facile à obtenir à partir des considérations suivantes. Séparons dans l'espace des vitesses une région s'étendant d'une certaine valeur $v_x < v_F$ à $v_x + dv_x$ (fig. 6.2). La sphère de Fermi découpe dans cette région un disque de volume $\pi(v_F^2 - v_x^2)\,dv_x$. Le nombre d'états contenus dans ce volume est précisément le nombre d'électrons dont la vitesse est comprise dans l'intervalle de v_x à $v_x + dv_x$:

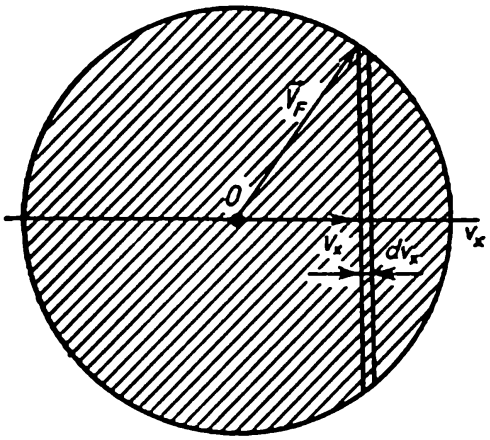

Fig. 6.2. Calcul de la répartition de Fermi des électrons suivant les vitesses.

$$n(v_x)\,dv_x = \frac{2m_n^3}{h^3}\pi(v_F^2 - v_x^2)\,dv_x. \qquad (6.8)$$

Cette fonction est représentée par la courbe supérieure de la figure 6.1, *b*. A $T > 0$ K, une petite partie des électrons ayant des vitesses v_x un peu inférieures à v_F acquièrent des vitesses un peu supérieures à v_F et la courbe de $n(v_x)$ se trouve modifiée comme l'indique la courbe en traits interrompus de la même figure.

Comme il fallait s'y attendre, les répartitions des électrons, aussi bien d'un gaz dégénéré que d'un gaz non dégénéré, sont symétriques par rapport à $v = 0$ de sorte que leur vitesse moyenne à l'état d'équilibre thermique est toujours nulle ($\sum \mathbf{v}_i = 0$).

L'état d'équilibre est établi et maintenu par suite de l'échange d'énergie et d'impulsions entre les électrons et le réseau cristallin. (Dans le cas de fortes concentrations des électrons c'est aussi l'interaction des porteurs de charges libres l'un avec l'autre qui commence à jouer un rôle important.) Les processus conduisant à un changement de l'énergie et de l'impulsion des électrons lors de leur interaction avec le réseau cristallin (ou l'un avec l'autre) sont appelés processus *de diffusion*.

La diffusion ne se produit que sur des *défauts* qui déforment le potentiel périodique régulier du cristal. Comme il a été dit au chapitre 4, les états des électrons dans un réseau parfait sont stationnaires et les électrons doivent y séjourner indéfiniment. (L'influence

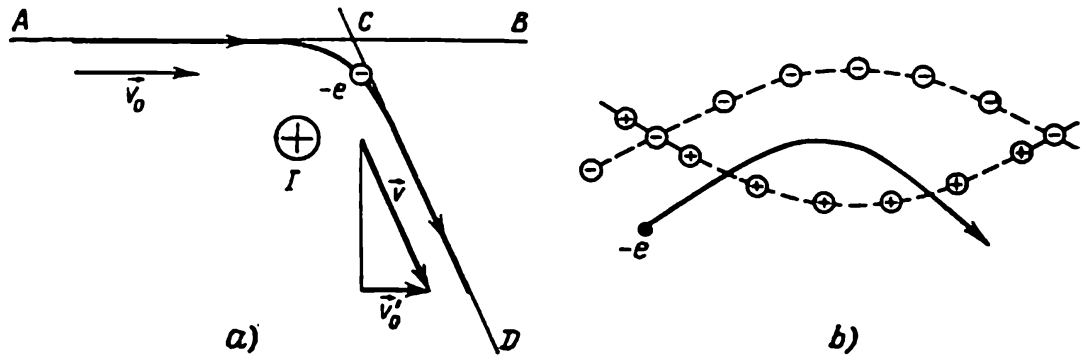

Fig. 6.3. Schémas illustrant la diffusion des électrons e par un atome d'impureté ionisé I (a) et par des vibrations optiques du réseau d'un cristal polaire (b) :
AB — direction de mouvement initial de l'électron ; v_0 — vitesse de ce mouvement; CD — direction de mouvement de l'électron après sa diffusion ; v — vitesse de ce mouvement ; v_0' — composante de la vitesse dans la direction du mouvement initial.

d'un réseau parfait sur le comportement d'un électron est prise en compte par introduction de la notion de masse effective.) Les défauts principaux ayant pour effet de provoquer la diffusion des électrons dans les cristaux sont les *atomes d'impureté ionisés* et les *vibrations thermiques du réseau*.

La diffusion sur les ions d'impureté peut être représentée classiquement comme une déviation d'un porteur de charge libre dans le champ de l'ion près duquel il passe (fig. 6.3, *a*). Seulement on ne devra pas oublier que cet ion doit être étranger pour que le champ qu'il engendre autour de lui diffère réellement du champ des ions propres du réseau par suite d'une autre valeur de la charge de cet ion ou d'autres dimensions et d'une autre structure des couches électroniques de l'ion d'impureté.

Quant à la diffusion sur les vibrations thermiques du cristal, ce phénomène semble au premier abord bien mystérieux. En effet, pourquoi l'électron ne se diffuse-t-il pas sur les ions immobiles ou sur les atomes de cristal mais se diffuse sur des ions vibrants bien que la vitesse de ces particules lourdes vibrantes soit toujours infé-

rieure de quelques ordres de grandeur à celle des électrons? Le fait est que l'apparition d'une onde élastique dans le cristal s'accompagne toujours de la création d'un champ électrique supplémentaire. On le voit le plus clairement sur l'exemple de vibrations optiques transversales des cristaux ioniques (fig. 6.3, *b*) ou sur l'exemple d'onde longitudinale dans le métal lorsque la concentration des ions devient augmentée dans la région de compression et diminuée dans la région de dépression. Mais dans les cristaux non polaires eux aussi les ondes élastiques donnent lieu à l'apparition de champs électriques supplémentaires, bien qu'ici c'est un effet du deuxième ordre de petitesse et il est difficile de représenter son origine de façon parlante. Dans tous les cas c'est précisément sur ces champs électriques supplémentaires que se produit la diffusion des électrons. Dans une telle diffusion l'énergie de l'électron varie de $\pm\hbar\omega$, où ω est la fréquence des vibrations du réseau, c'est-à-dire d'une quantité égale à l'énergie du phonon. En d'autres termes, il se produit soit une *absorption* soit une *émission d'un phonon* par un électron. L'impulsion de l'électron subit une variation de $\pm\hbar\mathbf{k}_{phon}$, où $\mathbf{k}_{phon}$ est le vecteur d'onde de la vibration du réseau ($\hbar\mathbf{k}_{phon}$ étant la quasi-impulsion du phonon).

§ 6.2. Dérive des porteurs de charges libres dans un champ électrique

L'approche la plus simple à la détermination du comportement des porteurs de charges libres dans un champ électrique *extérieur* consiste à les considérer comme des billes animées d'une agitation thermique désordonnée à laquelle se superpose, après l'application du champ, un mouvement orienté de l'ensemble — la *dérive*.

L'agitation thermique désordonnée des particules est caractérisée par une vitesse moyenne de mouvement $\langle v\rangle$, la distance moyenne parcourue par la particule entre deux collisions consécutives, c'est-à-dire par la longueur de libre parcours moyen $\langle\lambda\rangle$ et le temps de libre parcours moyen $\langle\tau\rangle$. Ces grandeurs sont liées entre elles par la relation évidente

$$\langle\tau\rangle = \langle\lambda\rangle/\langle v\rangle. \tag{6.9}$$

Le champ électrique extérieur communique aux particules chargées libres une accélération dirigée dans le sens du champ pour les trous et dans le sens opposé pour les électrons. Toutefois, par suite des collisions des particules (diffusion), leur mouvement redevient désordonné.

Sous l'action d'un champ d'intensité $\mathscr{E}$ une particule de masse effective m^* et de charge e se déplace avec une accélération

$$a = \frac{e}{m^*}\mathscr{E}.$$

Si l'on considère qu'après chaque collision la particule peut se mouvoir avec une même probabilité dans toute direction, c'est-à-dire que le mouvement des particules redevient tout à fait désordonné au bout du temps de libre parcours moyen $\langle\tau\rangle$, la vitesse moyenne de mouvement orienté (vitesse de dérive) est égale à

$$v_{der} = \frac{a\langle\tau\rangle}{2} = \frac{1}{2}\frac{e}{m^*}\langle\tau\rangle\,\mathscr{E}. \tag{6.10}$$

Ce qui vient d'être dit est illustré par la figure 6.4, où on a porté en ordonnées seulement la vitesse additionnelle (par rapport à la vitesse thermique) de mouvement de la particule sous l'effet du champ extérieur.

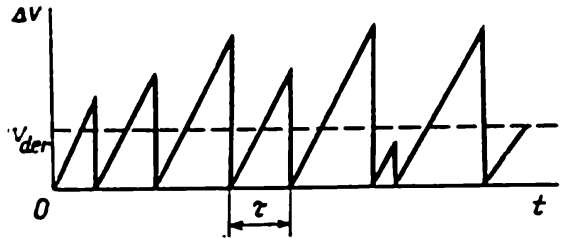

Fig. 6.4. Explication du calcul de la vitesse de dérive.

En réalité, la description des processus de diffusion est pourtant beaucoup plus complexe. Tout d'abord, si même on considère les microparticules comme des billes chargées, il est impossible d'établir avec exactitude le fait de diffusion d'un électron ou d'un trou et donc de déterminer $\langle\tau\rangle$ et $\langle\lambda\rangle$. Par exemple un électron qui passe près d'un ion peut changer la direction de son mouvement de 180°, 90°, 1° ou moins et il n'est pas clair quelle est une déviation arbitrairement petite par rapport à la direction de mouvement initiale qu'on peut considérer comme un témoignage de la diffusion de la particule. Il faut alors introduire la probabilité de diffusion dans les limites d'un angle solide déterminé et cette probabilité est différente pour les divers mécanismes de diffusion, les différentes directions d'angles de diffusion, etc. Il est donc clair que si la vitesse de dérive est déterminée par une expression de la forme (6.10), celle-ci contiendra une certaine grandeur effective $\langle\tau\rangle$ qui ne pourra être appelée temps de libre parcours moyen que conventionnellement.

Temps de relaxation de la fonction de répartition. La dérive des électrons dans un champ électrique extérieur, ainsi que d'autres phénomènes cinétiques peuvent être étudiés d'une façon plus rigoureuse à l'aide d'une *équation cinétique* dite *de Boltzmann*. Cette équation décrit les variations de la fonction de répartition des électrons suivant les vitesses sous l'effet d'un champ électrique extérieur et des processus de diffusion. L'apparition de la dérive des électrons sous l'action d'un champ électrique extérieur d'intensité $\mathscr{E}$ a pour effet de provoquer un déplacement de leur fonction de répartition suivant les vitesses de la quantité v_{der} dans le sens de la force **F** appliquée aux électrons (fig. 6.5). Après la suppression du champ, la fonction de répartition revient, par suite de la diffusion, à l'état d'équilibre : il se produit un processus dit *de relaxation*. On admet d'ordinaire que le processus de relaxation est décrit par l'équation

$$\Delta n\,(v_x) = \Delta n\,(v_x)\,|_{t_0} e^{-(t-t_0)/\tau_{rel}},$$

où t_0 est l'instant d'élimination du champ, τ_{rel} le temps de relaxation. Pendant le temps τ_{rel} l'écart de la fonction de répartition par rapport à la valeur d'équilibre et la vitesse de dérive diminuent de e fois. De la solution de l'équation cinétique de Boltzmann il résulte en particulier que

$$v_{der} = \frac{e}{m^*} \tau_{rel} \mathscr{E}. \qquad (6.11)$$

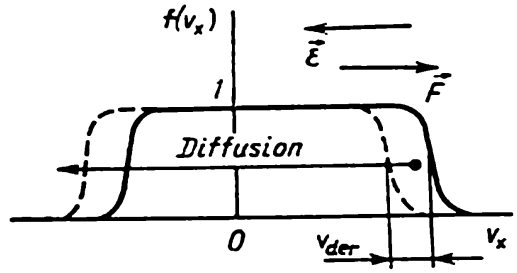

Fig. 6.5. Déplacement de la fonction de répartition des électrons suivant les vitesses ($v_y = v_z = 0$) dans un champ électrique extérieur pour un gaz dégénéré.

En comparant (6.11) et (6.10) nous voyons que τ_{rel} joue le même rôle que le temps de libre parcours moyen $\langle\tau\rangle$ en théorie simplifiée. C'est pourquoi, nous omettrons désormais l'indice affectant $\langle\tau\rangle$ dans (6.11) et, comme on le fait souvent, nous appellerons cette grandeur aussi bien temps de relaxation que temps de libre parcours moyen.

Mobilité des porteurs de charges. La grandeur égale au rapport de la vitesse de dérive à l'intensité du champ électrique est appelée *mobilité* des porteurs :

$$u = \frac{|v_{der}|}{\mathscr{E}} = \frac{e\langle\tau\rangle}{m^*}. \qquad (6.12)$$

En d'autres termes, la mobilité c'est la vitesse de dérive acquise par les électrons dans un champ d'intensité unité.

L'évaluation montre que pour les métaux tels que le cuivre $\tau \approx 2 \cdot 10^{-14}$ s, la mobilité de l'électron calculée par la formule (6.12) $u \approx 3 \cdot 10^{-3}$ m²/(V·s), la vitesse de dérive dans des champs d'intensité habituelle ($\mathscr{E} = 10^{-2}$ V/m) est $v_{der} \approx 0{,}3$ m/s. Cette quantité est inférieure de plusieurs ordres de grandeur à la vitesse de mouvement des électrons en l'absence du champ ($v_F = 1{,}6 \cdot 10^6$ m/s) de sorte qu'après l'application du champ électrique la vitesse moyenne des électrons dans un conducteur reste pratiquement inchangée.

Conductivité électrique. Comme l'agitation thermique désordonnée des électrons en état d'équilibre thermique ne peut pas conduire à leur mouvement orienté, on peut négliger la vitesse sommaire des électrons lors du calcul de la densité de courant électrique et de la conductivité électrique d'un conducteur et ne prendre en compte que la vitesse de dérive v_{der} acquise sous l'effet du champ $\mathscr{E}$.

Construisons, par la pensée, dans la profondeur d'un conducteur un cylindre de base égale à l'unité et de génératrice égale à v_{der} et dirigée le long de la dérive (fig. 6.6). Tous les électrons contenus dans ce cylindre passeront pendant 1 s par la base du cylindre et

constitueront un courant de densité

$$\mathbf{i} = -en\mathbf{v}_{der} = eu_n n\mathcal{E}, \tag{6.13}$$

où n est la concentration des électrons dans le conducteur. Suivant la loi d'Ohm $\mathbf{i} = \sigma\mathcal{E}$. On en déduit

$$\sigma = enu_n. \tag{6.14}$$

Pour les trous

$$\sigma = epu_p. \tag{6.14a}$$

En portant dans (6.14) la valeur de u donnée par (6.12), on obtient

$$\sigma = \frac{ne^2\langle\tau\rangle}{m_n} = \frac{ne^2}{m_n}\frac{\langle\lambda\rangle}{\langle v_n\rangle}. \tag{6.15}$$

Cette formule exprime la conductivité électrique d'un conducteur en fonction des paramètres microscopiques qui caractérisent son

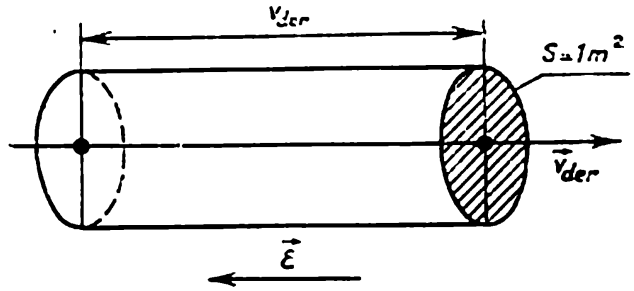

Fig. 6.6. Etablissement de l'expression pour la densité de courant.

gaz électronique. Le temps de relaxation τ_{rel} dépend dans le cas général de l'énergie des électrons. C'est pourquoi il doit être moyenné sur la collectivité des électrons et différemment pour les gaz dégénéré et non dégénéré.

Dans le cas d'un gaz *non dégénéré*, la densité de remplissage de la bande de conduction par les électrons est si faible que le comportement des électrons ne peut pas être limité par le principe de Pauli. Les électrons sont complètement libres en ce sens que le mouvement de chacun d'eux s'effectue sans subir une influence notable de la part des autres. Tous les électrons de conduction d'un gaz non dégénéré participent, indépendamment les uns des autres, à la création d'un courant électrique et à la conductibilité électrique d'un conducteur. Aussi, l'expression donnant la conductivité électrique d'un gaz non dégénéré doit-elle faire intervenir le temps moyen de relaxation $\langle\tau_{rel}\rangle$ de tous les électrons libres, obtenu en moyennant τ_{rel} sur toute la collectivité. Ceci étant, les expressions pour la mobilité et la conductivité électrique d'un gaz non dégénéré (par exemple

électronique) doivent s'écrire sous la forme suivante:

$$u = \frac{e}{m_n} \langle \tau(E) \rangle = \frac{e}{m_n} \frac{\langle \lambda(E) \rangle}{\langle v_n \rangle}, \tag{6.16}$$

$$\sigma = \frac{ne^2}{m_n} \langle \tau(E) \rangle = \frac{ne^2}{m_n} \frac{\langle \lambda(E) \rangle}{\langle v_n \rangle}. \tag{6.17}$$

Le calcul des moyennes se fait par des règles bien déterminées que nous n'examinons pas ici.

Il n'en est plus de même pour un gaz *dégénéré*. Dans un tel gaz la grande masse des électrons, dont l'énergie est inférieure à l'énergie de Fermi d'au moins 0,1 eV, ne peut pas participer aux processus de diffusion. Le fait est que de tels électrons peuvent se diffuser en passant aux états libres seulement si leur énergie augmente en même temps aux dépens de l'énergie des particules avec lesquelles ils entrent en collision. Mais les phonons possèdent dans le cristal une énergie qui ne dépasse pas quelques centièmes d'électron-volt et la diffusion sur les ions d'impureté s'effectue pratiquement sans changement de l'énergie de l'électron. C'est la raison pour laquelle seuls les électrons dont l'énergie est proche de celle de Fermi (*électrons de Fermi*) peuvent prendre part à la diffusion.

Sous l'action d'un champ électrique extérieur, tous les électrons du métal changent leur vecteur d'onde dans le sens de la force appliquée [suivant (4.17)] et la répartition des électrons est déplacée dans l'espace des v comme l'indique la figure 6.5. Toutefois, plus grand est ce déplacement, plus grand est le nombre des électrons de Fermi qui sont transférés de droite à gauche par suite du processus de diffusion, en tendant à rétablir la répartition d'équilibre des électrons suivant les impulsions, symétrique par rapport à l'origine des coordonnées. Par suite de l'action conjuguée du champ extérieur et des processus de diffusion sur les électrons, *toute* la sphère de Fermi se trouve déplacée de

$$\mathbf{v}_{der} = u\mathscr{E} = \frac{e\tau_F}{m_n} \mathscr{E} = \frac{e}{m_n} \frac{\lambda_F}{v_F} \mathscr{E}. \tag{6.18}$$

La valeur de v_{der} se détermine déjà non par les valeurs moyennes $\langle \tau \rangle$ ou $\langle \lambda \rangle$ et $\langle v \rangle$, comme c'est le cas d'un gaz non dégénéré, mais par le temps de relaxation τ_F ou la longueur de libre parcours moyen λ_F et la vitesse de déplacement v_F des électrons de Fermi. Du point de vue du calcul quantitatif de la conductivité seul le fait lui-même de déplacement de la répartition des électrons suivant les vitesses sous l'effet d'un champ qui importe et non pas son mécanisme physique. Aussi la conductivité électrique d'un gaz dégénéré peut-elle s'écrire comme précédemment sous la forme

$$\sigma = \frac{ne^2}{m_n} \tau_F = \frac{ne^2}{m_n} \frac{\lambda_F}{v_F}, \tag{6.19}$$

où n est la concentration totale des électrons dans la bande de conduction.

§ 6.3. Mobilité des porteurs de charges libres. Variation de la mobilité avec la température

Considérons maintenant la variation de la mobilité des porteurs de charges en fonction de la température. Cette étude sera faite séparément pour les hautes températures et les basses températures.

Dans le domaine des *hautes* températures c'est la diffusion des électrons sur les vibrations thermiques du réseau, c'est-à-dire la diffusion sur les *phonons*, qui a l'importance principale. Dans les cristaux non polaires la diffusion s'effectue de préférence sur des phonons acoustiques dont la concentration est plus forte que celle des phonons optiques. Comme le montre l'étude quantique, la longueur de libre parcours moyen ne dépend pas dans ce cas de l'énergie de l'électron ou du trou diffusé mais dépend uniquement de la concentration des phonons: $\lambda \sim 1/n_{phon}$. Vu que suivant le § 3.5 dans le domaine des hautes températures $n_{phon} \sim T$, on a $\lambda \sim 1/T$. En portant cette valeur de λ dans (6.16) et (6.18), on obtient les expressions suivantes pour la mobilité des électrons:

$$u \sim \langle\lambda\rangle/\langle v\rangle \sim T^{-1}/T^{1/2} = T^{-3/2} \text{ (gaz non dégénéré)}, \tag{6.20}$$

$$u \sim \lambda_{phon}/v_{phon} \sim T^{-1} \text{ (gaz dégénéré)}, \tag{6.21}$$

parce que v_{phon} ne dépend pratiquement pas de la température et pour un gaz non dégénéré $\langle v\rangle \sim T^{1/2}$.

Ainsi, *dans le domaine des hautes températures où la diffusion sur les vibrations thermiques du réseau est prédominante, la mobilité des porteurs constituant un gaz non dégénéré d'électrons (de trous) est inversement proportionnelle à $T^{3/2}$, la mobilité des porteurs dans un gaz dégénéré est inversement proportionnelle à T.*

Dans le domaine des *basses* températures c'est la diffusion sur des atomes d'impureté ionisés qui présente souvent l'importance principale. Le processus de diffusion consiste en ce que les ions d'impureté font dévier les électrons passant près d'eux, en modifiant en direction la vitesse de leur mouvement (v. fig. 6.3, *a*). Cette diffusion est pratiquement aussi désordonnée que la diffusion sur les vibrations du réseau parce que les électrons passent à de différentes distances des centres chargés et dévient d'angles différents.

Le problème de la diffusion des particules chargées par des centres chargés a été résolu pour la première fois par Rutherford lors de l'étude de la diffusion des particules alpha par des noyaux des éléments chimiques. La formule obtenue par Rutherford pour la longueur de libre parcours moyen λ se présente relativement à notre cas de la diffusion des électrons sur les ions d'impureté sous la forme suivante:

$$\lambda \sim \frac{1}{N_i} v^4 \left(\frac{\varepsilon}{Ze}\right)^2 m_n, \tag{6.22}$$

où N_i est la concentration des ions; v la vitesse des électrons; ε la permittivité du cristal; Ze la charge de l'ion diffusant.

Qualitativement, ce résultat est clair: plus grande est la vitesse des porteurs de charge, plus grande est leur masse effective et plus fort est l'affaiblissement du champ de l'ion dans le cristal (plus élevé est ε), plus faible est la déviation du porteur par rapport à la direction de mouvement initiale, plus longue est la distance qu'ils doivent parcourir pour que la composante de la vitesse de leur mouvement dans cette direction s'annule; lorsque la charge Ze de l'ion diffusant et la concentration N_i des ions augmentent, la valeur de λ doit naturellement diminuer.

En portant (6.22) dans (6.16) et (6.18), on obtient

$$u \sim \langle\lambda\rangle/\langle v\rangle \sim \langle v\rangle^3 \sim T^{3/2} \text{ (gaz non dégénéré)}, \tag{6.23}$$

$$u \sim \lambda_{phon}/v_{phon} \sim v_{phon}^3 = \text{const}\,(T) \text{ (gaz dégénéré)}. \tag{6.24}$$

On voit que *la mobilité des porteurs de charge dans le domaine des basses températures, due à la diffusion sur des impuretés ionisées, est proportionnelle à $T^{3/2}$ pour des conducteurs contenant un gaz électronique non dégénéré et ne dépend pas de T pour des conducteurs dont le gaz électronique est dégénéré.*

Dans le cas où plusieurs mécanismes de diffusion agissent *à la fois*, la mobilité des porteurs peut être déterminée en partant des considérations suivantes. La grandeur $w = 1/\tau$ est le nombre moyen de collisions subies par l'électron par unité de temps. Lorsque plusieurs mécanismes de diffusion indépendants agissent simultanément, le nombre total de collisions par unité de temps est égal à la somme des nombres de collisions dues aux différents mécanismes de diffusion: électron-phonon w_{phon}, électron-impureté w_{imp}, etc.:

$$w = w_{phon} + w_{imp} + \ldots. \tag{6.25}$$

Comme la mobilité est, suivant (6.12), proportionnelle au temps de libre parcours moyen τ, c'est-à-dire inversement proportionnelle au nombre w, la relation (6.25) peut être mise sous la forme

$$\frac{1}{u} = \frac{1}{u_{phon}} + \frac{1}{u_{imp}} + \ldots, \tag{6.26}$$

où u_{phon} et u_{imp} sont les mobilités dues à la diffusion uniquement sur les phonons et uniquement sur les impuretés respectivement. De la relation (6.26) il résulte que le mécanisme de diffusion décisif pour la détermination de la mobilité est celui qui conduit à la mobilité la plus basse.

Par exemple, si $u_{phon} = 1$ m²/(V·s) et $u_{imp} = 10^{-2}$ m²/(V·s), alors dans l'expression $1/u = 1/1 + 1/10^{-2}$ on peut négliger le premier terme et considérer que $u \approx u_{imp} = 10^{-2}$ m²/(V·s).

La figure 6.7, *a* montre les courbes théoriques de variation de la mobilité en fonction de la température dans un semiconducteur à gaz

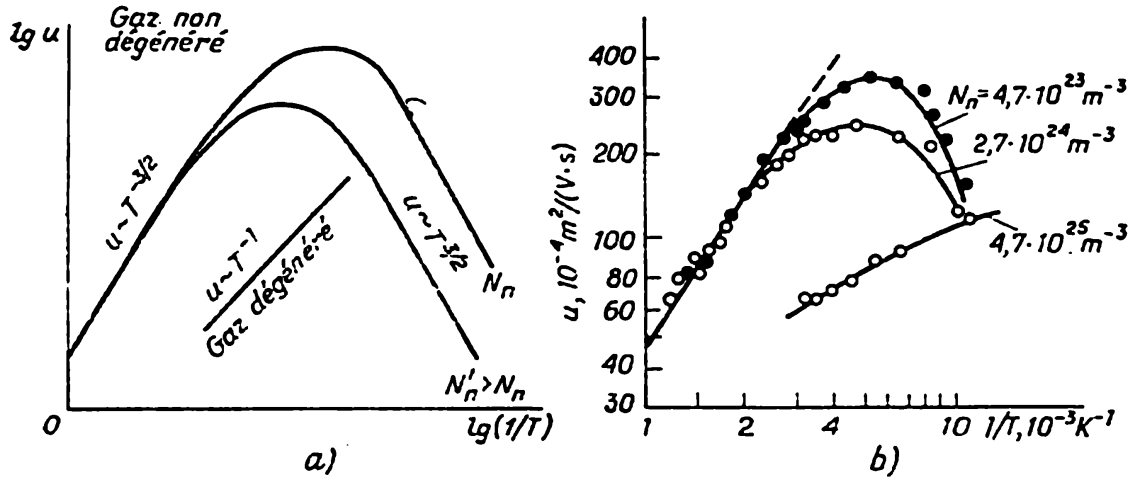

Fig. 6.7. Courbes théoriques (*a*) et expérimentales (*b*) de variation de la mobilité en fonction de la température dans un semiconducteur.

électronique non dégénéré et dégénéré et la figure 6.7, *b* les courbes expérimentales correspondantes pour le silicium contenant des quantités différentes N_n d'impureté dopante. Comme on le voit sur la figure 6.7, l'expérience confirme, pour l'essentiel, les conclusions de la théorie développée plus haut.

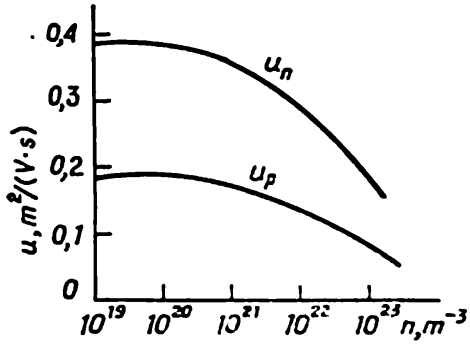

Fig. 6.8. Courbes expérimentales de variation de la mobilité des électrons u_n et des trous u_p en fonction de la concentration de l'impureté dopante dans le germanium.

Examinons maintenant la variation de la mobilité des porteurs dans un semiconducteur en fonction du degré de son dopage à une température fixée, par exemple ordinaire. Dans le cas d'un faible niveau de dopage la mobilité doit se déterminer par la diffusion sur les vibrations du réseau et donc doit être indépendante du dopage. Dans les semiconducteurs fortement dopés, elle doit se déterminer par la diffusion sur l'impureté de dopage ionisée et donc diminuer lorsque la concentration de l'impureté augmente. La figure 6.8 montre les courbes expérimentales de variation de la mobilité des porteurs dans le germanium en fonction de la concentration de l'impureté de dopage.

§ 6.4. Conductibilité électrique des semiconducteurs

Dans le cas général la conductivité électrique d'un semiconducteur se détermine aussi bien par les électrons (n) que par les trous (p):

$$\sigma = \sigma_n + \sigma_p = enu_n + epu_p. \qquad (6.27)$$

Toutefois dans le domaine des températures inférieures à la température de passage à la conductibilité intrinsèque T_i (v. fig. 5.8, c) la concentration des porteurs de charge d'un signe (porteurs majoritaires) est beaucoup plus forte que celle des porteurs de l'autre signe (porteurs minoritaires) si bien qu'il est logique de ne garder dans l'expression (6.27) que le terme correspondant aux porteurs de courant majoritaires.

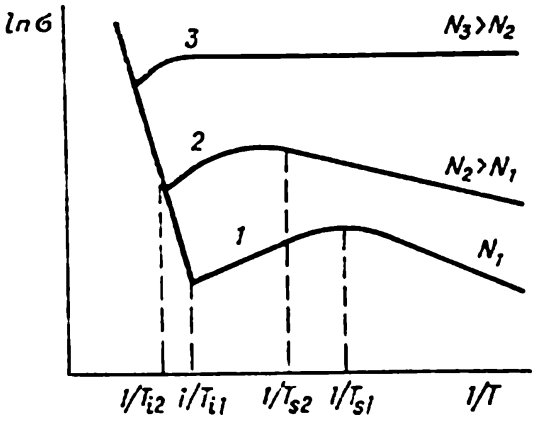

Fig. 6.9. Courbes de variation de la conductivité électrique avec la température des semiconducteurs de degré différent de dopage N (schématiquement).

Sauf le domaine de la température de déplétion de l'impureté, la concentration des porteurs de charge libres varie brusquement — exponentiellement — en fonction de la température. Par rapport à cette variation exponentielle la variation de la mobilité avec la température suivant une loi de puissance paraît assez faible et ne modifie pas dans l'ensemble la forme de la dépendance de ln σ par rapport à $1/T$ en comparaison de la dépendance de ln n par rapport à $1/T$. Les coefficients angulaires des dépendances correspondantes restent eux aussi pratiquement inchangés. Le mécanisme de diffusion prédominant aussi bien dans le domaine de la conductibilité intrinsèque que dans le domaine de très basses températures (inférieures à T_s) est la diffusion sur les phonons. Aux basses températures cela s'explique par une brusque diminution de la concentration de l'impureté ionisée.

Dans le domaine de déplétion de l'impureté (de T_s à T_i) la conductivité électrique diminue lorsque la température s'élève dans le cas des semiconducteurs faiblement dopés parce que le mécanisme de diffusion principal est la diffusion sur les phonons (courbe *1* de la figure 6.9). Pour un semiconducteur plus fortement dopé c'est la diffusion sur les impuretés qui peut prédominer au début mais, lorsque la température augmente, c'est la diffusion sur les vibrations du réseau qui devient prépondérante (courbe *2* de la figure 6.9).

Pour un semiconducteur fortement dopé, dont le gaz électronique est dégénéré presque jusqu'à la température de passage à la conducti-

bilité intrinsèque T_i, la mobilité ne dépend pas de la température. C'est seulement au voisinage de T_i où la dégénérescence du gaz électronique est levée et le rôle principal dans la diffusion est joué par les phonons, que la mobilité des électrons commence à diminuer lorsque la température augmente et l'exposant figurant dans sa dépendance par rapport à la température varie progressivement de zéro à —3/2. Dans un tel semiconducteur la concentration des porteurs de courant majoritaires est indépendante de la température jusqu'à T_i et toute la variation de σ est due à la variation de la mobilité. Ce cas est illustré par la courbe *3* de la figure 6.9.

§ 6.5. Conductibilité électrique des métaux

Comme la concentration n du gaz électronique dans les métaux est *indépendante* de la température, la variation de la conductivité électrique σ avec la température se détermine complètement par la variation, en fonction de la température, de la mobilité u des électrons constituant le gaz électronique dégénéré. Dans un métal suffisamment pur, la concentration d'impuretés est peu élevée et la mobilité de porteurs se détermine, jusqu'à des températures assez basses, par la diffusion des électrons sur les vibrations du réseau cristallin.

Dans le domaine des hautes températures, la mobilité des électrons est, suivant (6.21), inversement proportionnelle à T. En portant (6.21) dans (6.14), on obtient

$$\sigma \sim 1/T. \tag{6.28}$$

La résistivité est

$$\rho = 1/\sigma \sim T. \tag{6.29}$$

A des températures inférieures à la température de Debye cette dépendance cesse d'être valable.

Dans le domaine des *basses* températures la concentration du gaz phononique est, suivant le § 3.5, proportionnelle à T^3. Aussi le libre parcours moyen des électrons dû à leur diffusion sur les phonons doit-il être inversement proportionnel à T^3:

$$\lambda \sim T^{-3}.$$

Toutefois cette relation ne tient pas compte d'une circonstance importante suivante. Aux basses températures, l'impulsion moyenne des phonons est si petite devant l'impulsion des électrons de Fermi dans le métal que pour annuler le mouvement d'un électron dans une direction donnée il faut ν actes de sa diffusion sur les phonons et non pas un seul. Les calculs montrent que $\nu \sim T^{-2}$. Aussi le libre parcours moyen effectif λ_{eff} de l'électron sur lequel son mouvement devient désordonné a-t-il pour valeur

$$\lambda_{eff} = \nu\lambda \sim T^{-5}. \tag{6.30}$$

En portant (6.30) dans (6.19), on obtient

$$\sigma \sim T^{-5}, \quad \rho \sim T^{5}. \tag{6.31}$$

Notons que lors de la diffusion des électrons sur les phonons dans les semiconducteurs seuls les phonons de faibles impulsions participent à cette diffusion parce que les impulsions des électrons eux-mêmes sont relativement petites. Or, la concentration de tels phonons

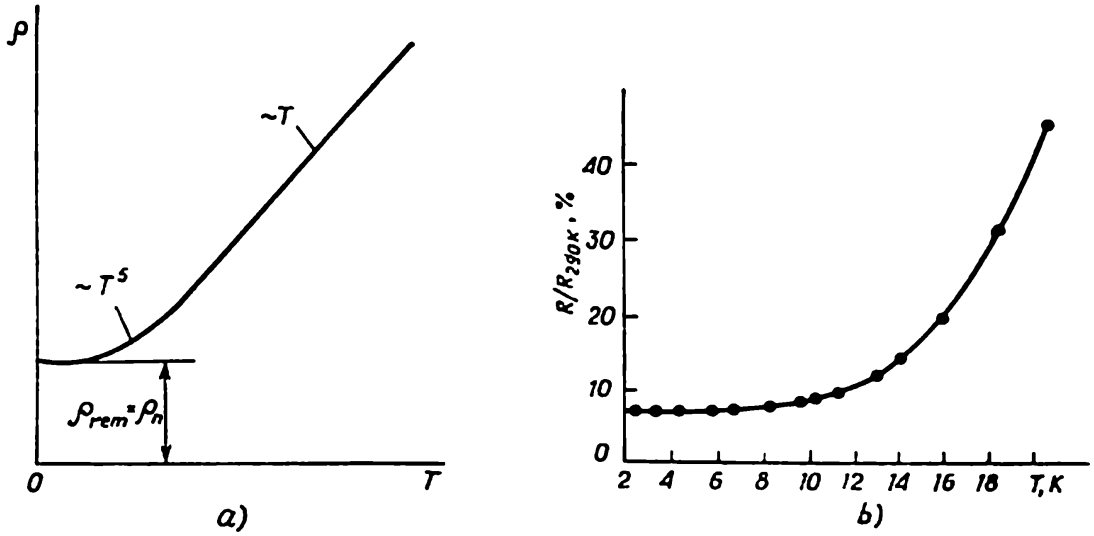

Fig. 6.10. Variation de la résistivité d'un métal pur avec la température: *a* — courbe théorique; *b* — variation de la résistance relative R/R_{290K} pour le sodium.

est également proportionnelle à la température même à des températures très inférieures à la température de Debye.

A des températures proches de 0 K, la concentration de phonons devient si faible que c'est la diffusion sur des atomes d'impureté contenus toujours en petite quantité même dans les métaux les plus purs, qui commence à jouer le rôle principal. En vertu de (6.24) la mobilité des électrons est dans ce cas indépendante de la température et la résistivité ρ prend une valeur constante que l'on appelle *résistivité rémanente* ρ_{rem}.

La figure 6.10, *a* montre la courbe schématique de variation de la résistivité des métaux purs en fonction de la température qui est confirmée assez bien par l'expérience (fig. 6.10, *b*).

Comme $\rho \sim 1/u$, à une température différente de 0 K, à la résistivité rémanente $\rho_{rem} = \rho_n$ est ajoutée, suivant (6.26), la résistivité ρ_{phon} due à la diffusion sur les vibrations thermiques du réseau. La résistivité totale du conducteur est dans ce cas égale à

$$\rho = \rho_n + \rho_{phon}. \tag{6.32}$$

Cette relation exprime la *règle de Matthiessen* bien connue concernant l'additivité de la résistance.

§ 6.6. Phénomène de supraconductivité

En 1911, en étudiant la variation de la résistivité du mercure en fonction de la température dans le domaine des basses températures, le physicien hollandais Kamerlingh-Onnes a découvert un phénomène étonnant: à une température de 4,2 K la résistance électrique du mercure disparaissait soudainement (fig. 6.11). Le mercure a cessé de présenter une résistance au passage d'un courant électrique. Ce phénomène a reçu le nom de *supraconduction* ou de *supraconductivité*. A l'heure actuelle il est établi que le phénomène de supraconductivité peut être observé dans des centaines de substances — les métaux purs, les alliages métalliques et les composés intermétalliques et même certains semiconducteurs dégénérés. La température de passage à l'état supraconducteur est appelée *température critique de transition*. Pour tous les métaux purs et les alliages ainsi que pour les semiconducteurs purs et les composés semiconducteurs qui ont été étudiés, cette température ne dépassait pas quelques degrés Kelvin et pour certains de ces matériaux quelques dixièmes et même centièmes de degré. De nombreuses tentatives faites en vue d'obtenir des composés ayant une température critique élevée n'ont donné pendant longtemps aucun résultat encourageant: la température de transition maximale de 23,2 K a été obtenue pour le composé Nb_3Ge. Une percée dans un nouveau domaine — celui de supraconduction à température élevée — à été effectuée à la fin de 1986 par Bernolz et Muller qui ont réussi à augmenter la température critique presque de 2 fois (jusqu'à 30 à 35 K) en utilisant des céramiques spéciales. Après cela, le nombre des ouvrages consacrés à la supraconduction à température élevée croissait en avalanche et, d'après certaines données, on a atteint une température critique supérieure à 250 K. Toutefois les tentatives de donner une interprétation sensée des résultats expérimentaux bien nombreux et souvent contradictoires n'ont pas encore abouti à la mise sur pied d'une théorie satisfaisante de ce phénomène exceptionnel. Aussi, dans ce qui suit allons-nous baser notre exposé sur des idées qui existaient avant 1986.

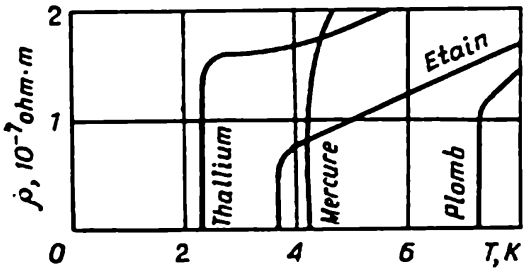

Fig. 6.11. Variation par saut de la résistivité des conducteurs lors du passage à l'état supraconducteur.

Soulignons une circonstance importante suivante. Comme il a été indiqué au § 6.5, si l'on réussissait à libérer les métaux des impuretés, alors au fur et à mesure qu'on s'approchait de 0 K, leur résistance devrait tomber progressivement jusqu'à zéro parce qu'un réseau exempt de défauts, présentant un potentiel rigoureusement

périodique, n'est pas en mesure de diffuser des porteurs de charge libres. Toutefois un tel comportement des métaux ne serait pas une supraconduction parce que premièrement, le passage d'une substance à l'état supraconducteur n'est pas lié en principe à la présence d'impuretés dans cette substance et deuxièmement, un tel passage se produit non pas d'une façon continue à mesure que la température baisse, mais soudainement à l'instant où la température de la substance devient égale à la température critique de transition T_c.

En outre, de nombreuses études ont montré que la supraconductivité ne se réduit seulement à l'annulation de la résistance électrique d'un conducteur (conductibilité parfaite). Une propriété non moins fondamentale de la substance à l'état supraconducteur est qu'elle montre un *parfait diamagnétisme*. Cette propriété découverte par Meissner et Ochsenfeld (1933) consiste en ce qu'en passant à l'état supraconducteur, une substance placée dans un champ magnétique (fig. 6.12, *a*) ne « congèle» pas le champ magnétique qu'elle contient, comme cela devrait se produire lors d'un simple passage de la substance à l'état de résistance nulle, mais le chasse dehors (fig. 6.12, *b*), ce qui est propre aux diamagnétiques parfaits possédant une perméabilité magnétique nulle. Ce phénomène a reçu le nom d'*effet Meissner-Ochsenfeld*.

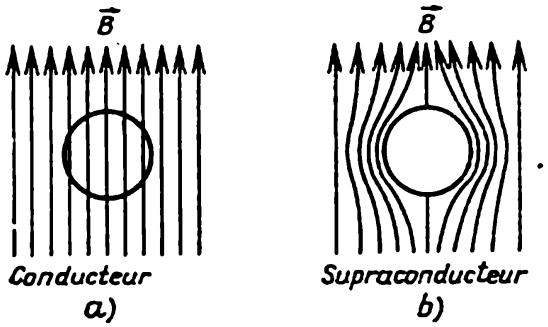

Fig. 6.12. Effet Meissner-Ochsenfeld (annulation du champ magnétique à l'intérieur d'un conducteur passant à l'état supraconducteur).

Ainsi, *l'essentiel de la supraconduction consiste en ce que la substance acquiert deux propriétés fondamentales non liées l'une à l'autre: une conductibilité parfaite et un diamagnétisme parfait.*

Nature de la supraconductivité. Malgré de grands efforts développés par de nombreux chercheurs pour étudier la nature de la supraconductivité, sa nature physique n'a été comprise qu'en 1957 lorsque Bardeen, Cooper et Schrieffer ont mis sur pied une théorie microscopique de ce phénomène qui a reçu plus tard le nom de *théorie de* BCS. Cette théorie est fondée sur l'idée qu'entre les électrons de conduction d'un métal peuvent s'exercer des forces d'*attraction* qui prennent naissance par suite de la *polarisation* du réseau cristallin par ces électrons.

Nous sommes accoutumés de considérer que deux électrons, deux particules portant des charges de même nom, ne peuvent subir qu'une répulsion coulombienne. Une telle situation, vraie pour des électrons se trouvant dans l'espace libre, peut changer de façon substantielle dans un cristal du fait que ce dernier contient d'autres

électrons et des ions chargés positivement qui compensent en moyenne les charges négatives des électrons. Dans de telles conditions, en plus des forces de répulsion entre les électrons peuvent être créées des forces d'attraction indirectes. En effet, un électron, qui se déplace dans le réseau, attire des ions chargés positivement, en les rapprochant un peu et en créant ainsi, le long de son parcours, une charge positive excédentaire du réseau polarisé qui peut attirer d'autres électrons. Cela est équivalent à l'apparition d'une force d'attraction entre les électrons qui n'agit pas directement mais par l'intermédiaire du réseau polarisé.

De ce point de vue, il faut s'attendre à ce que la supraconductivité se manifeste dans ceux des métaux qui se caractérisent par une forte interaction entre le gaz électronique et le réseau, qui conduit dans les conditions habituelles à une conductibilité électrique relativement basse. Et en effet, il se trouve que parmi les métaux purs les meilleurs supraconducteurs sont des métaux à haute résistivité — le plomb, le niobium, l'étain, le mercure et autres. En même temps, dans les métaux à haute conductivité tels que le cuivre et l'argent dont le gaz électronique se caractérise par une forte mobilité la supraconductivité n'est pas encore découverte.

La prédominance des forces d'attraction entre les électrons dans le réseau sur les forces de répulsion rend le gaz électronique dans les métaux instable à la formation, à partir des électrons, de *paires d'électrons* que l'on appelle *paires de Cooper* du nom de Cooper qui a montré le premier que la formation de telles paires est avantageuse du point de vue énergétique.

Les impulsions $\hbar\mathbf{k}$ des électrons constituant une paire sont à tout instant égales en module et de signes contraires, comme il se doit pour des particules identiques liées l'une à l'autre et en quelque sorte « tournant » l'une près de l'autre. Plus exactement, la fonction d'onde d'une paire est une superposition des états de $\mathbf{k}$ opposés, de valeur voisine des valeurs de Fermi k_F. Cette dernière assertion est due à ce que dans un gaz dégénéré de fermions ne peuvent interagir l'une avec l'autre que des particules près des niveaux énergétiques desquelles il existe des états libres. Cela est clair parce que toute interaction exige un changement d'état, ce qui est impossible si tous les états voisins sont occupés par d'autres particules.

L'attraction entre les électrons de Fermi a pour effet de réduire leur énergie totale. (Cette énergie doit être enlevée au gaz électronique lorsqu'il passe à l'état supraconducteur.) Dans le spectre énergétique des électrons il se forme alors une fente énergétique comme l'indiquent les figures 6.13, *a* et *b*. L'existence d'une telle fente peut être confirmée expérimentalement. Une de ces expériences consiste à étudier le spectre d'absorption par un supraconducteur des ondes électromagnétiques dans le domaine submillimétrique. A des températures voisines de 0 K l'absorption de l'énergie dans le supra-

conducteur ne se produit pas tant que l'énergie des quanta du rayonnement incident ne dépasse la largeur de la fente 2Δ, après quoi l'absorption croît rapidement jusqu'à sa valeur dans un métal normal parce que maintenant les électrons peuvent passer, en absorbant un phonon, dans des états énergétiques au-dessus de la fente. L'énergie 2Δ est dépensée pour la rupture de la paire (Δ pour chaque électron). Ces électrons deviennent indépendants l'un de l'autre, c'est-à-dire « normaux ».

Les paires peuvent être détruites aussi au détriment de l'énergie des vibrations thermiques du réseau. C'est pourquoi à $T > 0$ K il y a apparition des électrons dans des états au-dessus de la fente.

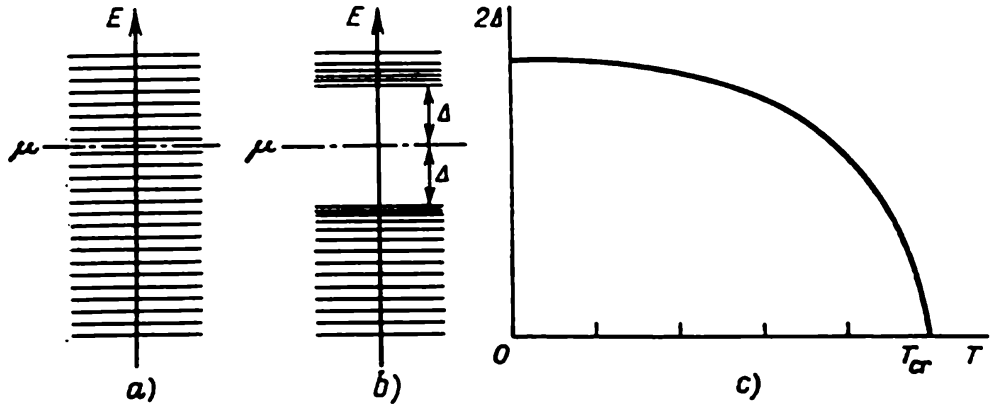

Fig. 6.13. Apparition d'un écart (une fente) dans le spectre énergétique d'un métal (*a*) lors de son passage à l'état supraconducteur (*b*) et variation de la largeur de l'écart énergétique en fonction de la température (*c*).

Il se produit simultanément un processus inverse de rétablissement des paires — le passage des électrons dans des états libres au-dessous de la fente. Ainsi, à chaque température, il s'établit une certaine concentration d'équilibre d'électrons « normaux ». La situation qui se présente ici est analogue à la génération thermique d'électrons et de trous dans un semiconducteur intrinsèque (v. § 4.7 et 5.4). Une différence substantielle réside pourtant en ce que, comme le montre la théorie et l'expérience, la largeur 2Δ de la fente énergétique dans un supraconducteur dépend du nombre des électrons non réunis en paires, en diminuant lorsque la concentration de ces derniers augmente et donc lorsque la température s'élève. La courbe de variation de la largeur de la fente énergétique en fonction de la température est représentée sur la figure 6.13, *c*.

La présence de la fente dans le spectre énergétique d'un supraconducteur explique pourquoi il peut être parcouru par des courants non amortis. En effet, revenons à l'examen de la figure 6.5 dont la flèche indique comment se déroule le processus de relaxation de la fonction de répartition des électrons dans un métal normal suivant les vitesses (la figure 6.14, *a* est une modification de cette figure).

Dans un métal supraconducteur de tels processus sont impossibles. Le fait est que pour faire varier l'impulsion d'un électron dans un métal normal aucune énergie peut ne pas être dépensée, l'impulsion peut changer seulement en direction. Quant au supraconducteur, toute variation de l'impulsion d'un seul électron provoquera inévitablement sa séparation de l'autre électron de la paire, c'est-à-dire la destruction de la paire, ce qui exige une énergie égale au moins à 2Δ. Aux basses températures, lorsqu'on observe la supraconductivité, le nombre de phonons d'énergie $\hbar\omega_{phon} > 2\Delta$ capables de détruire les paires (qui se rétablissent ensuite) est extrêmement petit.

Voyons maintenant pourquoi une paire ne peut pas être diffusée comme un tout (sans destruction). Revenons à cet effet à la figure 6.14, *b*. La variation supposée des impulsions des électrons de la paire y est représentée par des flèches *1* et *2*. Pour une telle variation des impulsions l'énergie cinétique totale des particules restera inchangée mais la paire dans son ensemble se déplace après la diffusion dans le sens opposé. Il est pourtant facile de voir qu'après une telle diffusion l'électron *2* doit arriver dans un état déjà occupé (les états occupés sont hachurés sur la figure), ce qui est interdit par le principe de Pauli. Ainsi, toute la collectivité des autres électrons empêche une telle diffusion de se produire et entretient un courant permanent.

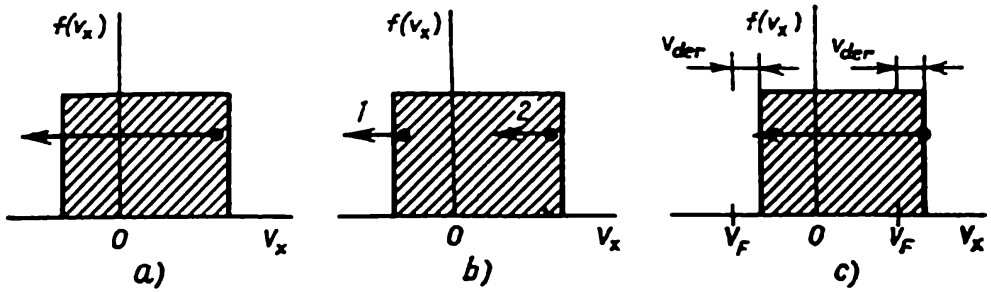

Fig. 6.14. Explication de la nature des courants non amortis dans un supraconducteur.

Enfin, considérons le processus de diffusion d'un électron de la paire à la suite duquel son énergie cinétique diminue (fig. 6.14, *c*) lorsqu'il est transféré de la partie de droite de la répartition dans la partie de gauche. Si cette diminution de l'énergie cinétique est suffisante pour la rupture de la paire, un tel processus est possible et l'état supraconducteur disparaît. Evaluons la vitesse de dérive à laquelle cela peut se produire. Comme la vitesse de l'électron considéré varie de $(v_F + v_{der})$ à $(v_F - v_{der})$, son énergie cinétique varie de

$$\Delta E_{cin} = m\,(v_F + v_{der})^2/2 - m\,(v_F - v_{der})^2/2 = \\ = 2mv_F v_{der}. \qquad (6.33)$$

La vitesse de dérive minimale $v_{der.cr}$ nécessaire à la destruction de la paire se détermine par la relation

$$2mv_F v_{der.cr} = 2\Delta, \tag{6.34}$$

et la densité de courant critique correspondant à $v_{der.cr}$ a pour valeur

$$i_{cr} = env_{der.cr} = en\Delta/(mv_F). \tag{6.35}$$

Pour des valeurs typiques de $n = 3 \cdot 10^{28}$ m^{-3}, $v_F = 10^6$ m/s et $2\Delta = 3$ meV, on obtient $i_{cr} = 10^{12}$ A/m^2, ce qui correspond à un courant de 10^6 A à travers un conducteur de 1 mm^2 de section. L'expérience montre toutefois qu'à travers le supraconducteur d'une telle section on n'arrive qu'à faire passer des courants de $\sim$100 A. Nous avons commis une erreur non dans l'évaluation de la densité de courant critique mais lors du passage de la densité de courant au courant lui-même quand nous avons admis implicitement que la répartition du courant suivant la section du supraconducteur était uniforme. Pour une telle répartition du courant le supraconducteur serait le siège d'un champ magnétique. Or, comme il a été dit plus haut, le supraconducteur est un diamagnétique parfait qui chasse le champ magnétique de son volume. Dans un tel cas, les courants ne doivent circuler que sur la surface du supraconducteur. La théorie prédit que ces courants passent dans une couche épaisse de $\lambda \approx 10$ à 100 nm. C'est à cette profondeur que le champ magnétique pénètre dans un supraconducteur en décroissant avec la distance x de la surface suivant la loi

$$B(x) = B(0) \exp(-x/\lambda), \tag{6.36}$$

où $B(0)$ est l'induction du champ près de la surface.

La supraconductivité se détruit lorsque la densité de courant à la surface $i(0)$ atteint sa valeur critique. Pour $\lambda = 35$ nm l'aire de la section de la couche superficielle d'un conducteur de 1 mm de diamètre, qui est parcourue par un courant, est égale à $\sim 10^{-4}$ mm^2. Pour une densité de courant critique $i_{cr} \approx 10^{12}$ A/m^2 un tel conducteur est en mesure de laisser passer un courant de $\sim$100 A, ce qui correspond à la réalité.

La supraconductivité peut être détruite aussi par un champ magnétique, ce qui découle directement de l'existence de i_{cr}. En effet, si un supraconducteur est placé dans un champ magnétique B, un courant non amorti est induit dans la couche superficielle et produit à l'intérieur du conducteur un champ B_{int} qui est dirigé dans le sens opposé à B et compense ce dernier. Lorsque B augmente, la densité de courant dans le supraconducteur et le champ compensateur B_{int} croissent. Toutefois, pour une certaine valeur de B, que l'on appelle *champ critique*, le courant induit dans le supraconducteur atteint sa valeur critique et la supraconductivité se détruit. Lorsque la température du supraconducteur s'élève, B_{cr} diminue. Suivant la théorie

de BCS cette diminution est décrite par la relation

$$B_{cr}(T) = B_{cr}(0)\,[1 - (T/T_{cr})^2], \qquad (6.37)$$

où $B_{cr}(0)$ est le champ critique à 0 K. La figure 6.15 montre les courbes expérimentales de $B_{cr}(T)$ pour le plomb et l'étain qui confirment cette dépendance. Il est visible sur cette figure qu'à des températures voisines de la température critique il suffit d'appliquer un champ magnétique faible pour détruire la supraconductivité, ce qui

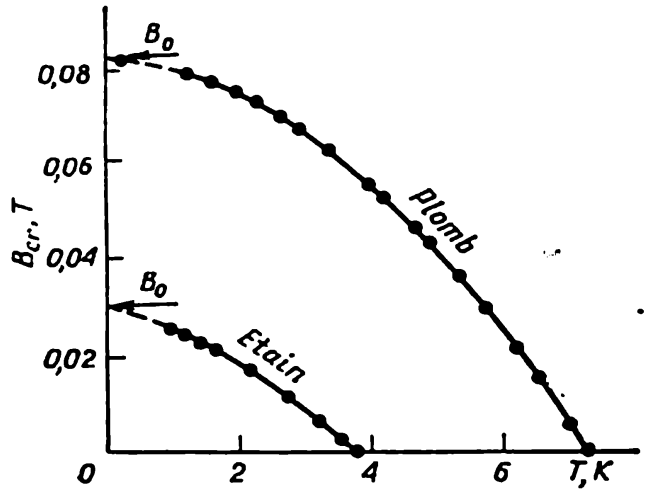

Fig. 6.15. Variation de l'induction d'un champ magnétique critique avec la température pour le plomb et l'étain. Les courbes en traits pleins correspondent à la formule (6.37).

est important pour la réalisation des dispositifs électroniques supraconducteurs.

La valeur de la grandeur λ, que l'on appelle *profondeur de pénétration*, dépend de la température et du matériau du supraconducteur. Dans des couches minces d'épaisseur inférieure à λ la valeur du champ critique B_{cr} détruisant la supraconductivité est d'un ordre de grandeur plus grande que dans des échantillons massifs. Néanmoins, l'existence d'un champ critique et de densités de courant critiques dans les supraconducteurs limite les possibilités de leur emploi pratique.

Un progrès considérable a été réalisé dans ce domaine quand on a obtenu des *supraconducteurs* dits *de deuxième espèce* ou « durs » représentés par de nombreux alliages et certains supraconducteurs à un seul constituant. La théorie des supraconducteurs de deuxième espèce a été développée par les physiciens soviétiques A. Abrikossov et L. Gorkov en 1957-1959. Les courants supraconducteurs I qui prennent naissance dans de tels matériaux, soumis à l'action d'un champ magnétique extérieur, circulent non seulement à la surface mais également dans la masse de conducteur (fig. 6.16).

Les régions cylindriques (filament de tourbillon) dans lesquelles circulent ces courants pénètrent tout le supraconducteur. Au centre de tels filaments les paires de Cooper n'existent pas et il n'y a pas de supraconductivité. C'est à travers ces filaments que le champ magnétique extérieur pénètre dans un supraconducteur de deuxième espèce.

A l'heure actuelle on a obtenu des supraconducteurs de deuxième espèce dont le champ critique dépasse 20 T (200 000 G) tandis que

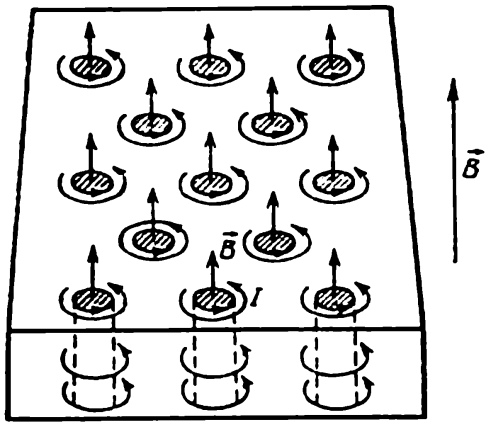

Fig. 6.16. Représentation schématique des filaments tourbillonnaires dans des supraconducteurs de deuxième espèce.

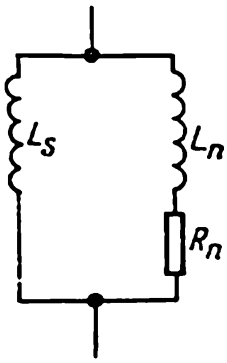

Fig. 6.17. Schéma équivalent d'un supraconducteur en courant alternatif.

dans les supraconducteurs de première espèce ce champ n'excède pas 0,1 T (1000 G).

Il est intéressant d'examiner le comportement des supraconducteurs dans un champ à haute fréquence. Comme chaque supraconducteur contient pratiquement toujours des électrons supraconducteurs (n_s) et des électrons normaux (n_n), le champ électrique variable accélère non seulement des paires de Cooper mais également des électrons normaux de sorte que le courant comporte une composante supraconductrice et une composante normale. Vu que les électrons, qu'ils soient supraconducteurs ou normaux, possèdent une masse, à cause de leur inertie le courant est déphasé en arrière sur l'intensité du champ à haute fréquence. Les paires de Cooper ne sont pas diffusées, c'est-à-dire que tout se passe comme si elles se déplaçaient dans un milieu sans frottement. Suivant la théorie classique, la vitesse des particules est dans ce cas en retard de $\pi/2$ sur la force périodique appliquée aux particules. C'est pourquoi la composante supraconductrice du courant de haute fréquence est déphasée en arrière de $\pi/2$ sur l'intensité du champ. Cela signifie que les paires de Cooper présentent une réactance pure. Quant aux électrons normaux, ils se diffusent en se déplaçant dans un milieu comme s'ils se déplaçaient dans un milieu avec frottement. Ils créent donc aussi bien une réactance qu'une résistance du supraconducteur. Cela permet de représenter un supraconducteur par un schéma équivalent de la figure 6.17.

La branche de gauche est constituée par l'inductance L_s des électrons supraconducteurs, la résistance de cette branche étant nulle. La branche de droite contient l'inductance L_n due à l'inertie des électrons normaux et la résistance R_n qui se crée par suite de la diffusion de ces électrons sur les défauts du réseau cristallin.

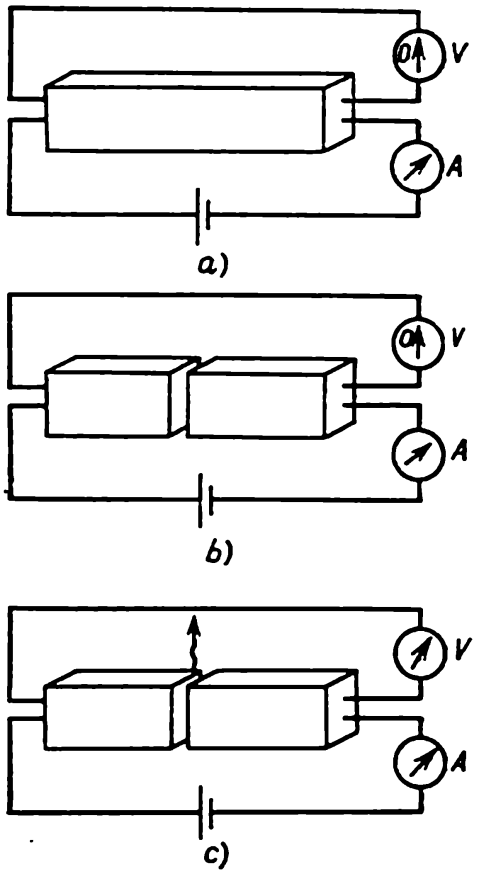

Fig. 6.18. Explication des effets Josephson.

Il convient de noter qu'à des fréquences $\nu \ll 2\Delta/h$ la résistance d'un supraconducteur est très grande par rapport à la réactance. En outre, elle augmente lorsque la température baisse parce que la concentration des électrons non réunis en paires diminue.

Effets Josephson. Comme il a été dit plus haut, la fonction d'onde d'une paire de Cooper est une superposition des états dont les vecteurs d'onde **k** sont opposés et ont une valeur voisine de celle de k_F. Toutes les fonctions d'onde des paires coïncident complètement. Une telle coïncidence est interdite pour les fermions par le principe de Pauli même pour deux particules mais les électrons qui se réunissent en une paire de Cooper ont non seulement des impulsions opposées mais également des spins opposés ; ils forment une nouvelle quasi-particule de spin total égal à zéro, c'est-à-dire un boson. Il est vrai que dans un gaz de bosons ordinaire il existe une répartition statistique des particules suivant les énergies. Pour ce qui est des paires de Cooper, elles se trouvent toutes dans un même état quantique et sont décrites par une seule fonction d'onde dans laquelle seule son amplitude dépend du nombre de particules. Une analyse théorique de cette circonstance a permis au physicien anglais D. Josephson de prédire en 1962 deux phénomènes intéressants qui ont reçu le nom d'*effets Josephson*. Actuellement, ces effets sont non seulement confirmés expérimentalement mais trouvent des applications pratiques notamment en microélectronique.

Etudions sommairement l'essentiel des effets Josephson. Branchons sur le supraconducteur une source de tension, un ampèremètre A et un voltmètre V (fig. 6.18, *a*). Lorsque le circuit est fermé,

le supraconducteur est parcouru par un courant I enregistré par l'ampèremètre. Comme la résistance du supraconducteur est nulle, la tension entre ses extrémités est elle aussi nulle. Divisons maintenant le supraconducteur en deux parties et plaçons entre elles un joint diélectrique d'épaisseur $d \approx 1$ nm, c'est-à-dire pratiquons dans le supraconducteur une fente transversale étroite. Comme l'a prédit Josephson, en mettant un tel supraconducteur en circuit, on peut observer un des deux effets suivants:

1. Le supraconducteur est parcouru toujours par un courant continu et la différence de potentiel entre ses extrémités est nulle. Autrement dit, le courant peut passer sans rencontrer aucune résistance non seulement à travers le supraconducteur mais aussi à travers sa fente si celle-ci est suffisamment étroite (fig. 6.18, *b*). Ce phénomène est connu sous le nom d'*effet Josephson continu*.

2. Une différence de potentiel continue V peut apparaître entre les extrémités du supraconducteur à fente. Dans ce cas, la fente rayonne une énergie électromagnétique à haute fréquence (fig. 6.18, *c*). Autrement dit, le circuit est parcouru dans ce cas non seulement par un courant continu mais aussi par un courant alternatif à haute fréquence. Ce phénomène est appelé *effet Josephson alternatif*.

Les effets Josephson sont dus au passage des paires d'électrons, par effet tunnel, d'un morceau de supraconducteur dans un autre à travers une fente étroite qui sépare ces morceaux. La théorie montre que le sens et l'intensité du courant par effet tunnel sont déterminés par la relation suivante:

$$I = I_0 \sin \varphi, \tag{6.38}$$

où φ est le déphasage entre les fonctions d'onde décrivant les paires de Cooper de part et d'autre de la barrière; I_0 l'intensité maximale du courant à travers la barrière dont la valeur est proportionnelle à la surface de la jonction tunnel et à la transparence de la barrière. Sa valeur peut être modifiée aussi par un champ magnétique dirigé le long du plan de la jonction Josephson perpendiculairement au flux de paires passant par effet tunnel d'un morceau du supraconducteur dans l'autre. Lorsque l'intensité H du champ magnétique augmente, le courant continu dû à l'effet Josephson subit des variations périodiques.

La relation (6.38) peut être expliquée à l'aide d'un modèle de deux pendules couplés par un faible ressort. Ce couplage a pour effet que lorsque les oscillations d'un des pendules sont en avance sur celles de l'autre, l'énergie est transférée du premier pendule au second. Le flux d'énergie atteint dans ce cas son maximum pour un déphasage de $\pi/2$. Si c'est le second pendule qui oscille en avance sur le premier, l'énergie est transférée du second pendule au premier.

Dans les contacts de Josephson les paires de Cooper passent d'un conducteur à l'autre. L'intensité et le sens du courant sont alors dé-

terminés par les mêmes relations de phase que pour des systèmes oscillants mécaniques faiblement couplés.

Lorsqu'on fait passer à travers la jonction de Josephson un courant I depuis une source extérieure, le déphasage φ varie automatiquement de façon à satisfaire à la condition (6.38). Dans ces conditions, l'intensité de courant I ne doit pas dépasser I_0. Le passage des courants de forte intensité peut donner naissance dans la jonction à une différence de potentiel V provoquant le passage par effet tunnel à travers la jonction des électrons non réunis en paires.

Lorsqu'il existe entre deux supraconducteurs une différence de potentiel, l'énergie des paires de Cooper de part et d'autre de la barrière diffère de $2\ eV$ (où $2e$ est la charge de la paire). Puisque l'énergie d'une particule est liée à la fréquence des ondes de de Broglie par la relation $E = h\nu$, il existe de part et d'autre de la jonction une différence de fréquence de de Broglie $\Delta\nu = 2eV/h$ et le déphasage φ des fonctions d'onde augmente de façon continue :

$$\varphi = 2\pi\Delta\nu t = 2\pi t\ (2eV/h). \tag{6.39}$$

En portant cette valeur de φ dans (6.38), on obtient l'expression suivante pour la composante supraconductrice du courant traversant la jonction :

$$I = I_0 \sin\left(\frac{4\pi eV}{h}\ t\right). \tag{6.40}$$

Maintenant c'est un courant alternatif et la fréquence de ses oscillations est égale à $2eV/h$. Pour $V = 1$ mV elle est de 485 GHz, ce qui correspond à une longueur d'onde $\lambda \approx 0{,}6$ mm.

§ 6.7. Effets galvanomagnétiques

Effet Hall. On appelle effet Hall un événement qui consiste en ce que le passage d'un courant électrique le long d'un ruban conducteur placé perpendiculairement à l'induction d'un champ magnétique extérieur fait naître dans ce ruban une différence de potentiel transversale par suite de l'interaction des porteurs de charge avec le champ magnétique.

Supposons qu'un conducteur ayant la forme d'une plaque rectangulaire est parcouru par un courant électrique I (fig. 6.19). En l'absence de champ magnétique, la différence de potentiel entre les points C et D situés sur l'une des surfaces équipotentielles est nulle. Si l'échantillon est placé dans un champ magnétique dont l'induction **B** est perpendiculaire à la direction du courant et au plan de l'échantillon, une différence de potentiel V_H dite *f.é.m. de Hall* apparaît entre les points C et D.

Comme le montre l'expérience, dans le cas des champs pas trop intenses; la tension V_H est proportionnelle à l'induction B du champ

et à l'intensité de courant I et inversement proportionnelle à l'épaisseur b de la plaque :

$$V_H = R_H BI/b = R_H Bia. \tag{6.41}$$

Ici, i est la densité de courant dans l'échantillon ; a la largeur de l'échantillon. Le coefficient de proportionnalité R_H est une constante caractérisant le matériau que l'on appelle *constante de Hall*,

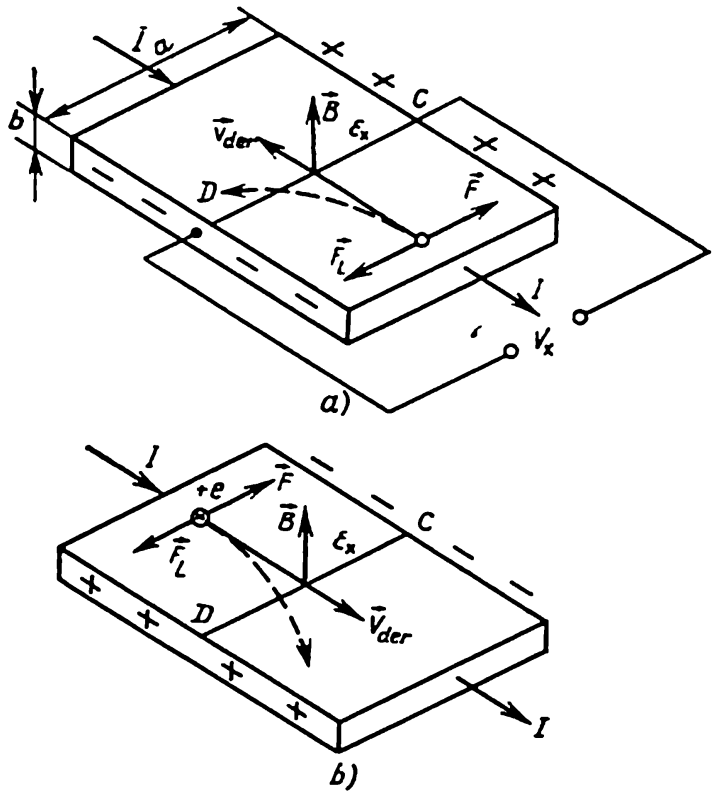

Fig. 6.19. Explication de l'effet Hall dans les semiconducteurs de type n (a) et de type p (b).

Elle a les dimensions L^3/Q (L étant la longueur et Q la charge électrique) et s'exprime en mètres cubes par coulomb.

Examinons la nature physique de l'effet Hall. Lorsque le conducteur est parcouru par un courant dans le sens indiqué sur la figure 6.19, a par une flèche, les électrons dérivent dans le sens opposé à la vitesse **v**. Chacun de ces électrons est soumis à la force de Lorentz exercée par le champ magnétique **B**

$$\mathbf{F}_L = -e\,[\mathbf{vB}], \tag{6.42}$$

où e est la charge de l'électron. Le sens d'action de cette force se détermine par la règle du tire-bouchon. L'angle fait entre **v** et **B** étant

égal à 90°, le module de la force de Lorentz a pour valeur

$$F_L = evB. \tag{6.43}$$

Sous l'action de la force de Lorentz les électrons dévient vers la frontière extérieure de la plaque, en la chargeant négativement (ligne en traits interrompus de la figure 6.19, *a*). Des charges positives non compensées s'accumulent sur la face opposée. Il en résulte l'apparition d'un champ électrique $\mathscr{E}_H = V_H/a$ dirigé de C vers D. Ici, V_H est la différence de potentiel entre les points C et D (*f.é.m. de Hall*).

Le champ $\mathscr{E}_H$ exerce sur les électrons une force $F = -e\mathscr{E}_H$ dirigée contre la force de Lorentz. Pour $F = F_L$ le champ électrique transversal équilibre la force de Lorentz et une accumulation ultérieure des charges électriques sur la face latérale de la plaque cesse. A partir de la condition d'équilibre

$$evB = -e\mathscr{E}_H \tag{6.44}$$

on tire

$$\mathscr{E}_H = vB. \tag{6.45}$$

En tenant compte que la densité de courant dans le conducteur $i = env$, où n est la concentration des électrons, on obtient $v = i/(en)$. En portant cette expression dans (6.45), on trouve

$$V_H = [1/(en)]\, Bia. \tag{6.46}$$

Ainsi, la théorie conduit à une expression pour V_H qui coïncide avec l'expression (6.4) établie expérimentalement. La constante de Hall a dans ce cas pour valeur

$$R_H = 1/(en). \tag{6.47}$$

Lors de l'établissement de la formule (6.45) nous n'avons tenu compte que de la vitesse de mouvement orienté (vitesse de dérive) des électrons. Cela est naturel parce que l'agitation thermique désordonnée des porteurs de charge ne peut pas conduire à leur déplacement orienté dans le champ magnétique. En outre, nous avons admis implicitement que tous les porteurs possèdent une même vitesse de dérive. Une telle hypothèse simplificatrice peut être justifiée pour les métaux et les semiconducteurs dégénérés dans lesquels le courant est transporté par les électrons qui possèdent pratiquement une même énergie (énergie de Fermi), mais elle est inapplicable aux semiconducteurs non dégénérés dans lesquels les porteurs de charges ont des énergies différentes et peuvent avoir des vitesses de dérive différentes, étant donné que leur mobilité dépend de la vitesse d'agitation thermique (plus exactement du temps de libre parcours moyen). Par exemple, lors de la diffusion par des impuretés chargées, la vitesse de dérive des porteurs de haute énergie (porteurs

possédant des vitesses élevées d'agitation thermique) est plus grande que celle des porteurs de basse énergie; par contre, lors de la diffusion par des vibrations thermiques du réseau, la vitesse de dérive des électrons de haute énergie est inférieure à celle des électrons de basse énergie. Une théorie plus rigoureuse, qui tient compte de cette circonstance, donne pour la constante de Hall l'expression suivante:

$$R_H = A/(en), \tag{6.48}$$

où A est une constante égale pour les réseaux atomiques à 1,17 si les porteurs sont diffusés sur les vibrations thermiques et à 1,95 s'ils sont diffusés sur les impuretés ionisées.

Pour les semiconducteurs du type p le sens de dérive des porteurs de charge coïncide avec le sens du courant. La force de Lorentz est dirigée dans ce cas lui aussi de C vers D parce que le signe de la charge et le sens de dérive des porteurs changent en même temps (fig. 6.19, *b*). Mais cette force s'exerce maintenant sur des particules chargées positivement de sorte que le point D se trouve porté à un potentiel positif par rapport au point C. Cela signifie que par le signe de la f.é.m. de Hall on peut déterminer le signe des porteurs de courant. On considère, par convention, que le signe de la constante de Hall est positif si le courant est transporté par les trous et négatif s'il est transporté par les électrons. Compte tenu de cette règle, on écrit la constante de Hall sous la forme suivante:

$$R_H = A/(ep) \tag{6.49}$$

pour les semiconducteurs du *type p* avec une concentration des trous égale à p et sous la forme

$$R_H = -A/(en) \tag{6.50}$$

pour les semiconducteurs du *type n* avec une concentration des électrons égale à n.

Pour les semiconducteurs à conduction *mixte*, lorsque les concentrations des électrons et des trous sont comparables l'une à l'autre, la constante de Hall est calculée à l'aide de la relation suivante:

$$R_H = \frac{A}{e}\,\frac{u_p^2 p - u_n^2 n}{(pu_p + nu_n)^2}. \tag{6.51}$$

Cette formule se réduit à (6.49) pour $p \gg n$ et à (6.50) pour $n \gg p$.

Pour les semiconducteurs intrinsèques, dans lesquels $n = p = n_i$, la formule (6.51) prend la forme

$$R_H = \frac{A}{en_i}\,\frac{u_p - u_n}{u_p + u_n}. \tag{6.52}$$

Vu que généralement la mobilité des électrons est plus grande que celle des trous, le signe de la constante de Hall est, en règle générale, négatif.

L'effet Hall est un puissant instrument d'étude expérimentale des propriétés des porteurs de charges dans les semiconducteurs. En mesurant la constante de Hall R_H, on peut calculer la concentration des porteurs de charges et déterminer leur signe d'après le sens de la f.é.m. de Hall.

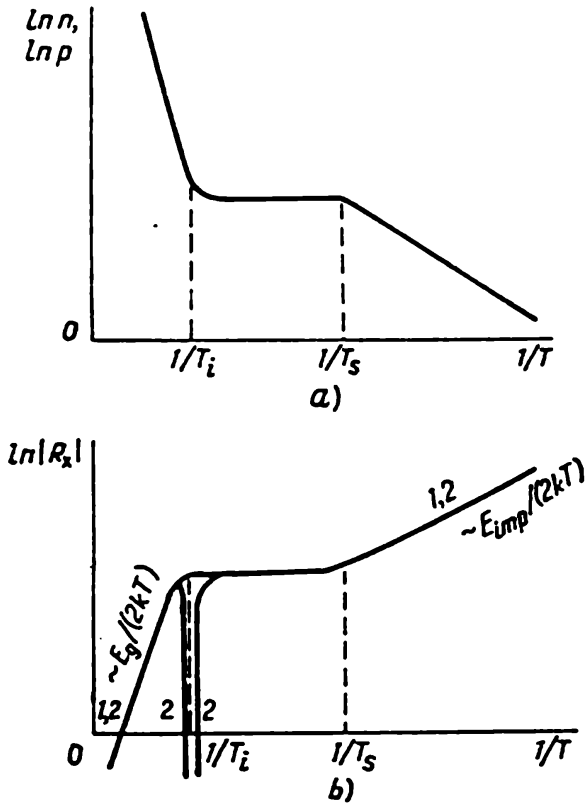

Fig. 6.20. Variation de la concentration des porteurs de charge (*a*) et de la constante de Hall (*b*) pour les semiconducteurs de type *n* (courbe *1*) et de type *p* (courbe *2*) en fonction de la température.

En multipliant la constante de Hall R_H par la conductivité électrique $\sigma = enu_n$, on obtient

$$R_H\sigma = Au_n. \qquad (6.53)$$

Connaissant A on peut calculer à l'aide de cette relation la mobilité u_n des porteurs de charges.

Ce sont les mesures de la variation de la constante de Hall avec la température qui présentent une importance particulière. Elles permettent d'établir la variation, en fonction de la température, de la concentration des porteurs de charges et, en association avec les mesures de la conductivité électrique, la variation de la mobilité en fonction de la température. En utilisant la variation, en fonction de la température, de la concentration des porteurs ou de la constante de Hall dans le domaine de la faible ionisation des impuretés, on peut déterminer l'énergie d'ionisation E_{imp} et, dans le domaine de la conduction intrinsèque, aussi la largeur de la bande interdite E_g. La figure 6.20 montre la variation, en fonction de la température, de la concentration des porteurs de charges (fig. 6.20, *a*) et de la constante de Hall (fig. 6.20, *b*) en coordonnées correspondantes, lorsque ces dépendances sont représentées par des droites. La courbe *2* de la figure 6.20, *b* se rapporte à un semiconducteur du type *p*. Dans le domaine de la conductibilité d'impureté la constante de Hall est dans un tel semiconducteur positive alors que dans le domaine de la conductibilité intrinsèque elle est négative. Lors du passage à la

conductibilité intrinsèque la constante de Hall R_H change de signe en passant par zéro et le logarithme de $|R_H|$ tend dans ce cas vers $-\infty$.

Il a été indiqué au début du paragraphe que la tension de Hall est proportionnelle à l'induction du champ magnétique, si l'intensité de ce champ n'est pas trop élevée. Les calculs montrent que le critère

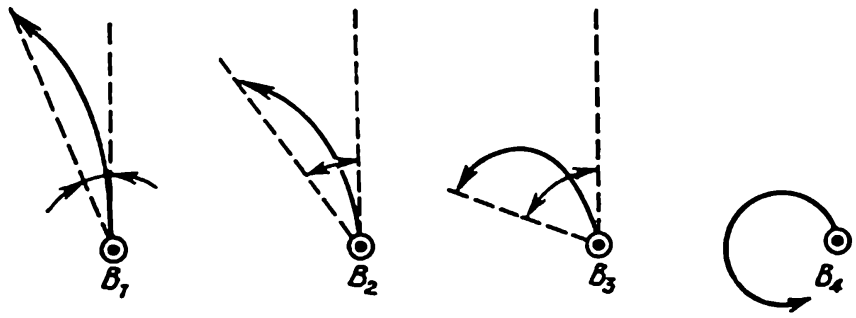

Fig. 6.21. Variation de la trajectoire des porteurs de charge dans des champs magnétiques de différente induction ($B_4 > B_3 > B_2 > B_1$).

de « faiblesse » du champ et donc d'applicabilité des expressions obtenues plus haut à la f.é.m. de Hall et à la constante de Hall est la condition suivante:

$$Bu \ll 2\pi, \tag{6.54}$$

où B est l'induction du champ; u la mobilité des porteurs de charges. Les évaluations correspondantes montrent que pour le germanium et le silicium dans lesquels les mobilités sont égales à 0,38 et 0,16 m²/(V·s) respectivement, les champs sont faibles, si leur induction $B < 16$ T et $B < 40$ T, tandis que pour InSb de $u_n = 8$ m²/(V·s) on peut considérer comme faibles des champs de $B < 0{,}8$ T. La cause qui fait varier la constante de Hall lorsque B augmente consiste en ce que dans les champs magnétiques très intenses les porteurs de charges non seulement s'écartent, pendant le temps de libre parcours moyen, de la direction de mouvement initiale mais encore, comme l'indique la figure 6.19, s'enroulent autour des lignes du champ magnétique (sur la figure 6.21 cela se produit pour un champ B_4).

La proportionnalité qui existe entre la f.é.m. de Hall et l'induction de champ magnétique (dans des champs peu intenses) permet d'utiliser des échantillons semiconducteurs comme capteurs pour la mesure de l'intensité de champs magnétiques. Les dimensions de tels capteurs peuvent être suffisamment petites (par exemple 0,5 × 0,5 mm), ce qui permet d'effectuer à leur aide des mesures avec une bonne résolution en espace.

L'inertie de l'effet Hall se détermine par le temps de relaxation de Maxwell, c'est-à-dire qu'elle est extrêmement faible. Cela permet d'utiliser les capteurs de Hall pour la mesure des champs magné-

tiques à haute fréquence, pour la détermination de l'intensité de courant par le champ magnétique engendré, etc.

Comme le montre la relation (6.41), la f.é.m. de Hall est proportionnelle au produit de l'intensité du courant parcourant le capteur par l'induction de champ magnétique. Cela permet d'utiliser l'effet Hall pour la multiplication de deux signaux, ce qui doit se faire par exemple dans les mesureurs de puissance, dans les détecteurs de phase, dans les calculateurs.

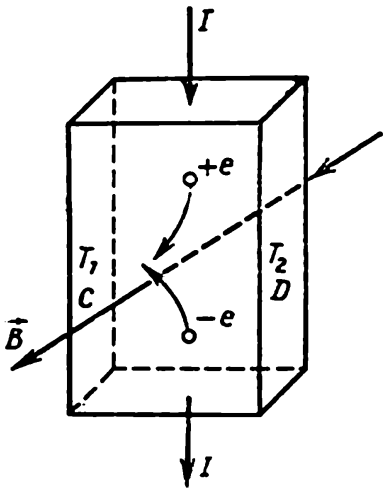

Fig. 6.22. Schéma illustrant l'apparition de l'effet Ettingshausen dans un semiconducteur intrinsèque.

Ces dernières années, le nombre d'applications des capteurs de Hall a connu une brusque croissance, en s'étendant à la radiotechnique et à l'électronique.

Effet Ettingshausen. L'effet Ettingshausen accompagne l'effet Hall et consiste en ce que le passage d'un courant I dans un conducteur placé dans un champ magnétique transversal **B** (fig. 6.22) dans une direction perpendiculaire au champ magnétique et au courant, fait apparaître un gradient de température. Cet effet est le plus marqué dans les semiconducteurs intrinsèques. Comme il a été indiqué plus haut, dans de tels semiconducteurs les électrons et les trous sont déviés par le champ magnétique dans un même sens (vers la face C de la figure 6.22). Il en résulte que sur l'une des faces de l'échantillon la concentration des électrons et des trous devient supérieure à celle d'équilibre et la recombinaison y l'emporte sur la génération thermique de porteurs de charges tandis que sur l'autre face (face D de la figure 6.22) la concentration des porteurs de charge est au contraire inférieure à celle d'équilibre et la génération thermique l'emporte sur la recombinaison. Ceci étant, la chaleur est dépensée pour la génération de paires électron-trou dans une partie de l'échantillon et se dégage, par suite de la recombinaison de ces paires, dans l'autre partie de l'échantillon — une différence de température $T_1 - T_2$ prend naissance dans l'échantillon (fig. 6.22).

L'effet Ettingshausen s'observe aussi dans le domaine de la conductibilité d'impureté des semiconducteurs. Dans ce cas, la cause de son apparition est la différence entre les vitesses de dérive des porteurs de charges possédant des vitesses différentes d'agitation thermique et, de ce fait, des temps différents de libre parcours moyen des porteurs « froids » et « chauds ». Suivant la formule (6.44), le champ de Hall $\mathscr{E}_H$ ne compense l'action de la force de Lorentz que

pour des porteurs de charges animés d'une vitesse de dérive moyenne. Les porteurs dérivant à une plus grande vitesse dévient vers la force de Lorentz alors que ceux qui dérivent à une plus faible vitesse dévient dans le sens opposé. C'est le mécanisme prédominant de la diffusion des porteurs qui détermine ceux parmi eux qui sont les plus « chauds ». Par exemple, si c'est la diffusion sur les vibrations thermiques du réseau qui est prédominante, les électrons « chauds » ont une mobilité inférieure à celle des électrons « froids » et, de ce fait, dévient vers la face D (fig. 6.19, *a*) de la plaque et cette face s'échauffe; la face C vers laquelle dévient les électrons « froids » se refroidit.

L'effet Ettingshausen peut être utilisé dans les installations de climatisation, de refroidissement, de thermostatisation et autres où on a à pomper de la chaleur.

Les effets galvanomagnétiques ne se limitent pas à ceux examinés plus haut. Toutefois ce sont les effets Hall et Ettingshausen qui ont reçu le plus d'applications pratiques. Les renseignements sur les autres effets peuvent être trouvés dans tout manuel de physique des semiconducteurs.

§ 6.8. Effets thermomagnétiques

Les effets thermomagnétiques interviennent lorsqu'un conducteur présentant un gradient de température est soumis à l'action d'un champ magnétique.

Soit par exemple un échantillon (fig. 6.23) présentant un gradient de température suivant sa longueur de la face supérieure B vers la face inférieure A ($T_B < T_A$). Ceci étant, des flux résultants de porteurs de charges libres se dirigent de la face A vers la face B. Ils sont dus à ce qu'à l'extrémité chaude de l'échantillon les vitesses moyennes d'agitation thermique désordonnée sont plus grandes qu'à l'extrémité froide (thermodiffusion) ainsi qu'à ce que la concentration des porteurs à l'extrémité chaude peut se trouver plus forte qu'à l'extrémité froide (diffusion). Le passage des porteurs de courant libres vers l'extrémité froide a pour effet de provoquer l'apparition d'une différence de potentiel entre les extrémités froide et chaude de l'échantillon que l'on appelle *force électromotrice thermo-électrique* (pour plus de détail, voir § 10.6). Cette différence de potentiel fait naître un flux opposé aux flux de thermodiffusion et de diffusion, à savoir le flux de dérive des porteurs. Par exemple, si l'échantillon

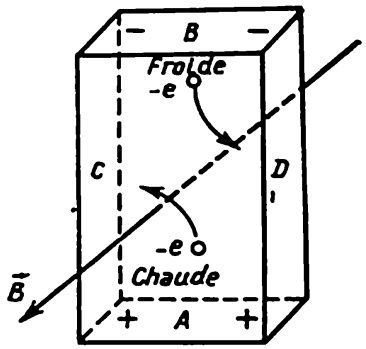

Fig. 6.23. Schéma illustrant l'apparition de l'effet Righi-Leduc.

est fait en semiconducteur de type n, la thermodiffusion et la diffusion des électrons ont pour résultat l'apparition d'une charge négative sur la face supérieure et d'une charge positive sur la face inférieure de l'échantillon (fig. 6.23). Un flux d'électrons dérive alors de la face supérieure vers la face inférieure. La f.é.m. thermo-électrique croît tant que le flux de dérive ne s'égalise avec le flux thermique opposé. Toutefois, ces deux flux sont transportés par des électrons dont les vitesses moyennes de l'agitation thermique sont différentes. La vitesse moyenne du mouvement désordonné des porteurs est plus élevée dans le flux thermique que dans le flux de dérive. Dans un champ magnétique transversal, ces flux dévient vers les faces opposées C et D. Il en résulte que la face C vers laquelle dévie le flux thermique sera plus chaude que la face D et il y aura apparition d'un gradient de température transversal proportionnel au gradient longitudinal. Ce phénomène est connu sous le nom d'*effet Righi-Leduc*.

Les vitesses moyennes de l'agitation thermique désordonnée des porteurs dans les flux longitudinaux opposés étant différentes, la déviation des porteurs par le champ magnétique est un peu différente de sorte que dans le sens transversal il se crée aussi une certaine différence de potentiel (*effet Nernst-Ettingshausen transversal*).

La courbure des trajectoires des électrons dans le champ magnétique conduit à ce que la f.é.m. thermo-électrique entre les faces A et B subit une certaine variation (*effet Nernst-Ettingshausen longitudinal*). Le coefficient de conductibilité thermique suivant la longueur de l'échantillon varie lui aussi (*effet Magee-Reghi-Leduc*).

CHAPITRE 7

PROPRIÉTÉS MAGNÉTIQUES DES SOLIDES

Les éléments et dispositifs fondés sur l'utilisation des phénomènes magnétiques qui s'observent dans les solides sont d'application courante en électronique des solides. Ce sont les amplificateurs paramagnétiques, les circuits magnétiques à ferrites en technique des hyperfréquences, les mémoirs magnétiques à ferrites, les films magnétiques minces et les mémoires à base de ces films, les cryotrons magnétiques, les intégrateurs magnétiques, les dispositifs fondés sur la commande de déplacement des domaines magnétiques, les aimants permanents pour les systèmes de prémagnétisation et beaucoup d'autres.
Le présent chapitre contient des renseignements principaux sur les propriétés magnétiques des solides et sur leur nature physique nécessaires pour comprendre les principes physiques de fonctionnement de tels dispositifs.

§ 7.1. Classification des solides suivant leurs propriétés magnétiques

Caractéristiques magnétiques des milieux. Plaçons un corps isotrope de volume V dans un champ magnétique uniforme d'intensité H et d'induction $B = \mu_0 H$ (μ_0 étant une *constante* appelée perméabilité magnétique du vide). Sous l'effet du champ le corps se magnétise en acquérant un moment magnétique M. Le quotient de ce moment par le volume du corps est appelé *aimantation* J_m du corps:

$$J_m = M/V. \tag{7.1}$$

L'aimantation est une grandeur vectorielle. Dans les corps magnétiques isotropes, le vecteur aimantation $\mathbf{J}_m$ est parallèle ou antiparallèle à $\mathbf{H}$. Dans le système SI, le moment magnétique s'exprime en ampères par mètre carré et l'aimantation en ampères par mètre. Le rapport de J_m à l'intensité de champ H porte le nom de *susceptibilité magnétique*:

$$\varkappa = J_m/H. \tag{7.2}$$

Un corps aimanté, placé dans un champ extérieur, produit un champ propre $B_i = \mu_0 J_m = \mu_0 \varkappa H = \varkappa B_0$ qui est, dans les corps magnétiques isotropes, parallèle ou antiparallèle au champ extérieur B_0. Le champ résultant dans les magnétiques a pour expression

$$B = B_0 + B_i = (1 + \varkappa) B_0 = \mu_0 (H + J_m). \tag{7.3}$$

La grandeur

$$\mu = 1 + \varkappa \tag{7.4}$$

est appelée *perméabilité magnétique relative*. En portant (7.4) dans (7.3), on obtient

$$B = \mu B_0 = \mu_0 \mu H. \tag{7.5}$$

Dans le système SI, la grandeur H s'exprime en ampères par mètre et l'induction en teslas.

Suivant la valeur et le signe de la susceptibilité magnétique, tous les corps peuvent être rangés en trois groupes : *diamagnétiques*, *paramagnétiques* et *ferromagnétiques*. Le tableau 7.1 indique les valeurs

Tableau 7.1

Diamagnétiques	$\varkappa \cdot 10^6$	Paramagnétiques	$\varkappa \cdot 10^6$	Ferromagnétiques	$\varkappa$
Bismuth	−18	CaO	580	Fer	1000
Cuivre	−0,9	$FeCl_2$	360	Nickel	240
Germanium	−0,8	$NiSO_4$	120	Cobalt	150
Silicium	−0,3	Platine	26		

de la susceptibilité magnétique de divers matériaux dans un champ extérieur de faible intensité à la température ordinaire.

Propriétés magnétiques des corps diamagnétiques et paramagnétiques. La susceptibilité magnétique des corps diamagnétiques $|\varkappa| \ll 1$, elle est négative et ne dépend ni de l'intensité du champ

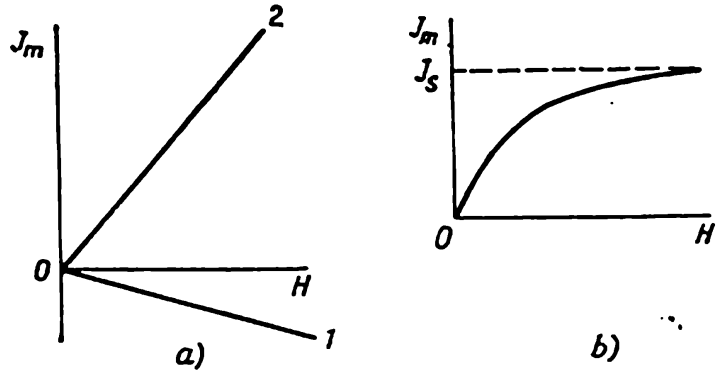

Fig. 7.1. Variation de l'aimantation J_m en fonction de l'intensité H du champ magnétique :

a — pour les diamagnétiques (*1*) et les paramagnétiques (*2*) dans des champs pas trop intenses et aux températures ordinaires et élevées ; *b* — pour les paramagnétiques aux basses températures et dans des champs intenses.

extérieur, ni de la température. Les corps diamagnétiques sont aimantés dans le sens contraire au champ de sorte qu'ils sont repoussés des régions où le champ est plus intense.

La susceptibilité magnétique des corps paramagnétiques est aussi $|\varkappa| \ll 1$ mais ses valeurs sont positives. Ces corps s'aimantent dans un sens parallèle au champ extérieur. La figure 7.1, *a* montre la

variation de l'aimantation J_m en fonction de H pour les diamagnétiques et les paramagnétiques. Dans les deux cas, $J_m \sim H$, ce qui signifie que $\varkappa$ est indépendant de H. Toutefois, pour les paramagnétiques une telle dépendance ne s'observe que dans des champs relativement faibles et aux températures élevées ; dans des champs intenses et aux basses températures la valeur de J_m (H) tend asymptotiquement vers la valeur limite J_s correspondant à la « saturation » magnétique correspondante des paramagnétiques (fig. 7.1, *b*). En

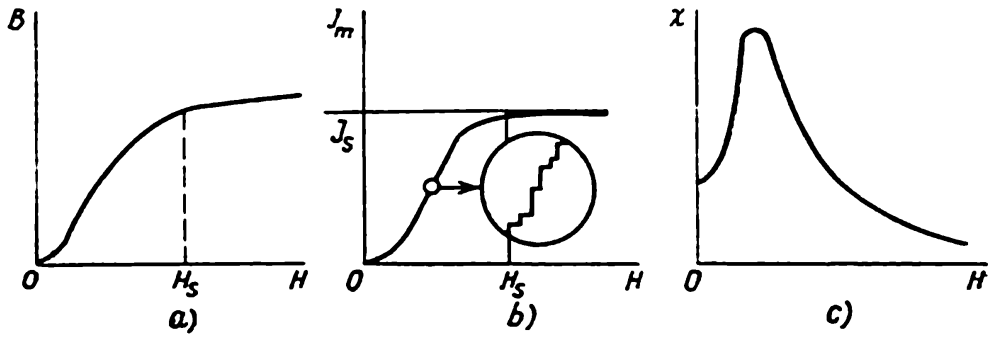

Fig. 7.2. Aimantation des corps ferromagnétiques. Variations en fonction de l'intensité H du champ magnétique :

a — de l'induction B ; *b* — de l'aimantation J_m (à droite on a représenté, à une plus grande échelle, un tronçon de la courbe d'aimantation) ; *c* — de la susceptibilité magnétique χ.

outre, la susceptibilité magnétique $\varkappa$ des corps paramagnétiques dépend de la température :

$$\varkappa = C/T. \tag{7.6}$$

Cette dépendance a été établie par Curie et s'appelle *loi de Curie*. La constante C est dite *de Curie*.

Propriétés magnétiques des corps ferromagnétiques. Chez les ferromagnétiques, dont le représentant typique est le fer, la susceptibilité $\varkappa$ est aussi positive mais sa valeur est beaucoup plus grande que celle des paramagnétiques et elle dépend de H. En plus du fer ce groupe comprend le nickel, le cobalt, le gadolinium et certains autres métaux ainsi que plusieurs alliages. Les lois régissant l'aimantation des ferromagnétiques ont été étudiées pour la première fois par le physicien russe Stolétov. La figure 7.2 montre la variation, en fonction de H, de l'induction B, de l'aimantation J_m et de la susceptibilité magnétique $\varkappa$ d'un fer doux. Lorsque l'intensité du champ magnétique augmente, l'induction B et l'aimantation J_m croissent d'abord rapidement, ensuite leur croissance se ralentit et, à partir d'une certaine valeur H_s, l'aimantation atteint une valeur pratiquement limite J_s. Stolétov a donné à cet état le nom de *saturation magnétique*. Quand l'état de saturation magnétique est atteint, l'induction $B = \mu_0 (H + J_m)$ croît proportionnellement à H.

La figure 7.3 montre la courbe représentant un cycle complet de magnétisation d'un ferromagnétique dans les deux directions. On voit sur cette figure que lorsque la direction de la magnétisation est changée, la variation de B est en retard sur la variation de H et il se trouve que pour $H = 0$ l'induction B ne s'annule pas mais est égale à B_r. Le phénomène d'un tel retard de B sur H est appelé *hystérésis magnétique* et l'induction B_r est dite *induction rémanente* ou *magnétisme rémanent*. Pour supprimer l'induction rémanente, il faut appliquer un champ démagnétisant H_c que l'on appelle *force coercitive*. La courbe fermée $AB_rH_cA'B'_rH'_cA$ qui décrit le cycle complet de magnétisation dans deux directions est appelée *cycle d'hystérésis*. L'aire du cycle d'hystérésis est proportionnelle au travail de renversement de l'aimantation d'un volume unité du ferromagnétique. Au cours du renversement de l'aimantation ce travail se transforme totalement en chaleur. C'est la raison pour laquelle les ferromagnétiques peuvent s'échauffer fortement dans des champs magnétiques alternatifs de fréquence élevée.

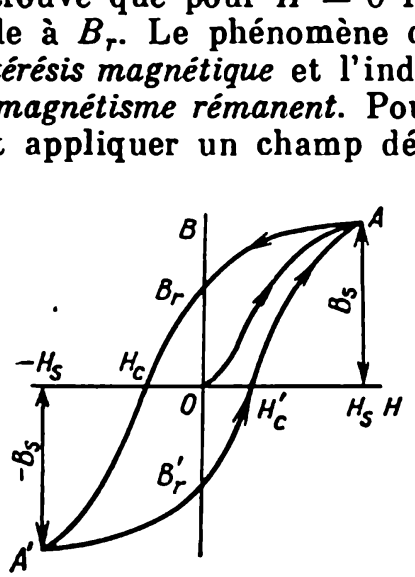

Fig. 7.3. Cycle d'hystérésis.

Suivant la forme et l'aire du cycle d'hystérésis, les ferromagnétiques sont divisés en *magnétiquement doux* et *magnétiquement durs*. Les matériaux magnétiquement doux possèdent une faible force coercitive et une haute perméabilité. Pour les meilleurs alliages de ce type H_c est égale à $\sim$0,3 A/m et μ atteint des valeurs de $\sim 10^5$. Les matériaux magnétiquement durs sont caractérisés par une haute force coercitive ($H_c \approx 10^4$ à 10^5 A/m) et, en règle générale, par un fort magnétisme rémanent ($B_r \approx 1{,}5$ T).

L'échauffement des corps ferromagnétiques fait varier leurs propriétés magnétiques : $\varkappa$, μ, J_m accusent une diminution. Pour chaque ferromagnétique il existe une température Θ_C à laquelle il perd ses propriétés ferromagnétiques. Cette température est appelée *point ferromagnétique de Curie*. Indiquons, à titre d'exemple, les valeurs des points de Curie (°C) pour le fer, le cobalt et le nickel :

Fer	770
Cobalt	1150
Nickel	360

A des températures supérieures à Θ_C, les ferromagnétiques deviennent paramagnétiques avec la variation linéaire de $1/\varkappa$ en fonction de la température T qui est caractéristique des paramagnétiques. Cette dépendance est exprimée de façon approchée par la *loi de*

Curie-Weiss:

$$\varkappa = C/(T - \Theta_C), \tag{7.7}$$

où C est la constante de Curie.

§ 7.2. Nature du diamagnétisme. Résonance cyclotronique

On sait que si l'axe d'une toupie tournant à grande vitesse ne coïncide pas avec la verticale, la toupie effectue, en plus de la rotation autour de son axe, encore une rotation avec l'axe autour de la

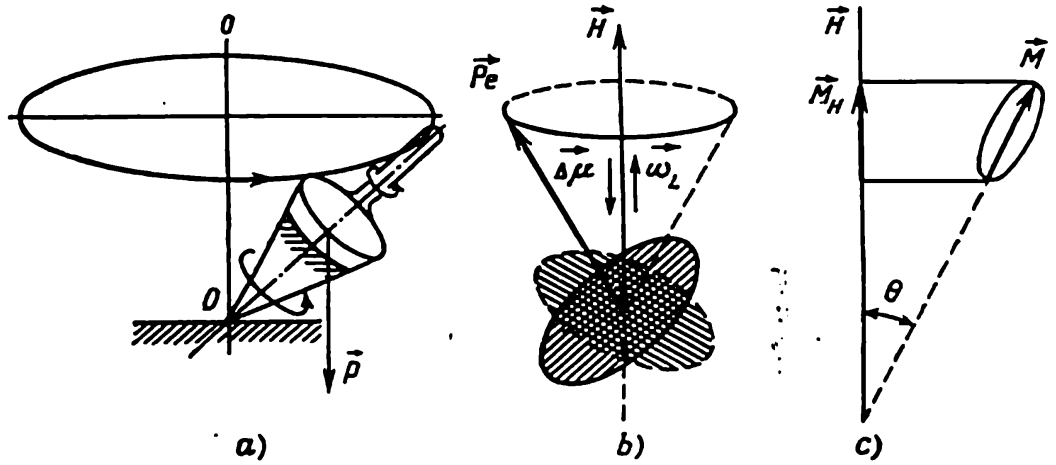

Fig. 7.4. Précession d'une toupie autour de la direction OO parallèle à la force de pesanteur **P** (*a*); précession d'une orbite électronique autour de la direction du champ magnétique **H** avec une vitesse angulaire $\vec{\omega}_L \downarrow\downarrow$ **H** conduisant à l'apparition d'un moment magnétique $\vec{\Delta\mu} \downarrow\uparrow$ **H** (*b*); moment magnétique **M** et sa projection M_H sur la direction du champ **H** (*c*).

verticale (fig. 7.4, *a*). Ce mouvement additionnel est appelé *précession*. Il s'effectue, sous l'action de la force de pesanteur **P**, autour de la direction OO parallèle à cette force.

Une situation analogue se présente dans le cas où un atome est placé dans un champ magnétique **H** (fig. 7.4, *b*). Un électron gravitant sur une orbite constitue un courant fermé qui possède un moment magnétique. Le champ magnétique tend à placer l'orbite de l'électron perpendiculairement à **H**, ce qui provoque sa précession autour de la direction du champ. Le calcul montre que la fréquence angulaire d'une telle précession est égale à

$$\omega_L = [e/(2m_e)]\, \mu_0 H,$$

où e, m_e sont respectivement la charge et la masse de l'électron. La fréquence ω_L est appelée *fréquence de Larmor*.

La précession d'une orbite électronique à la fréquence ω_L est équivalente à un mouvement additionnel de l'électron autour du champ **H** qui fait naître un courant fermé $\Delta I = e\nu_L = e\omega_L/(2\pi) =$

$= [e^2/(4\pi m)]\ \mu_0 H$ possédant un moment magnétique

$$\Delta\mu = \Delta I S = [\mu_0 e^2 S/(4\pi m)]\ H,$$

où S est la surface limitée par le contour que décrit l'électron par suite de sa précession autour de **H**. Comme la vitesse angulaire $\vec{\omega}_L$ est dirigée le long du champ **H**, le moment magnétique induit $\vec{\Delta\mu}$ est dirigé dans le sens opposé à **H**. C'est la naissance de ce moment qui détermine l'apparition du diamagnétisme.

Le diamagnétisme est inhérent à toutes les substances mais souvent il n'est pas manifeste sur le fond du paramagnétisme et du ferromagnétisme plus forts. Il ne se rencontre dans sa pureté que dans des substances dont le moment magnétique résultant des atomes est nul.

Les porteurs de charges libres dans les corps solides qui possèdent une composante non nulle de la vitesse, perpendiculaire au champ magnétique, décrivent des orbites circulaires (v. § 4.11) et engendrent un champ magnétique opposé au champ extérieur. Aussi, le gaz de porteurs libres possède-t-il un diamagnétisme qu'on appelle *diamagnétisme de porteurs libres* ou encore *diamagnétisme de Landau*. En ordre de grandeur il est comparable à celui des couches électroniques.

Comme il a été montré au § 4.10, le spectre énergétique des électrons et des trous varie sous l'effet d'un champ magnétique: il y a apparition de niveaux de Landau. Si une onde électromagnétique dont l'énergie des quanta $\hbar\omega$ est égale à la distance entre les niveaux de Landau voisins $\hbar\omega_c$ tombe sur un cristal, elle provoquera le passage des électrons à des niveaux plus élevés qui s'accompagne d'une forte absorption de l'énergie de rayonnement. Ce phénomène porte le nom de *résonance cyclotronique*. (Une faible crête d'absorption s'observe aussi pour $\hbar\omega = 2\hbar\omega_c$.) Comme $\hbar\omega_c = \hbar eB/m_n$, d'après la fréquence d'absorption par résonance on peut déterminer la masse effective des porteurs de charges libres. Etant donné que la diffusion des électrons conduit à un élargissement des niveaux d'énergie, la résonance cyclotronique ne s'observe nettement que dans des cristaux suffisamment parfaits, aux basses températures, et dans des champs magnétiques intenses ($B \gtrsim 1$ T) capables d'écarter des niveaux de Landau flous voisins à des distances pour lesquelles ils ne se recouvrent pas. Pour de tels champs ω_c se situe dans le domaine des hyperfréquences (pour $B = 1$ T, la fréquence $\omega_c = 2\cdot 10^{12}$ s^{-1}, ce qui correspond à une longueur d'onde $\lambda \approx 0{,}1$ cm). Comme le montre l'expérience, pour réaliser la résonance cyclotronique à de telles fréquences, il est plus commode de faire varier non pas la fréquence ω du rayonnement incident mais l'induction B du champ magnétique. La méthode fondée sur la résonance cyclotronique s'est avérée d'une grande fécondité et de puissance pour l'étude de la structure des bandes d'énergie des solides.

§ 7.3. Nature du paramagnétisme

Paramagnétisme des diélectriques. Le paramagnétisme est observé dans des substances dont les atomes possèdent un moment magnétique résultant non nul, c'est-à-dire représentent des dipôles magnétiques élémentaires dont l'interaction est négligeable. Comme exemple de telles substances on peut citer O_2 et NO_2, les solutions aqueuses et les hydrates cristallins de sels des éléments de transition du groupe du fer et des terres rares, ainsi que d'autres substances. Dans les solutions aqueuses et les hydrates cristallins, les ions d'éléments de transition possédant des moments magnétiques non compensés sont séparés l'un de l'autre par les molécules d'eau et interagissent l'un avec l'autre assez faiblement, ce qui est une condition nécessaire pour la manifestation du paramagnétisme.

Rappelons brièvement la nature des moments magnétiques des atomes. Chaque électron gravitant autour du noyau crée un courant fermé $I = e\nu$, où e est la charge de l'électron et ν la fréquence de sa révolution sur l'orbite. Le moment magnétique de ce courant est $M = IS = e\nu S$, où S est l'aire de l'orbite. Comme $S = \pi r^2$ et $\nu = v/(2\pi r)$, où v est la vitesse linéaire de mouvement de l'électron sur l'orbite, on a

$$M = \mu_l = ver/2. \tag{7.8}$$

Le moment magnétique de l'électron dû à son mouvement autour du noyau est appelé *moment magnétique orbital.* Le moment μ_l est dirigé perpendiculairement au plan de l'orbite dans le sens opposé au moment cinétique de l'impulsion ayant pour valeur

$$\mathcal{L}_l = m_e vr, \tag{7.9}$$

où m est la masse de l'électron. Le rapprochement entre (7.8) et (7.9) donne

$$\vec{\mu}_l = -\frac{e}{2m_e}\mathcal{L}_l.$$

Suivant les lois de la mécanique quantique $\mathcal{L}_l$ peut prendre seulement la série de valeurs suivantes: $\mathcal{L}_l = \hbar\sqrt{l(l+1)}$, où l est le *nombre quantique orbital.* De ce fait, il vient

$$\mu_l = -\frac{e\hbar}{2m_e}\sqrt{l(l+1)} = -\mu_B\sqrt{l(l+1)}, \tag{7.10}$$

où

$$\mu_B = e\hbar/(2m_e) = 9{,}27\cdot 10^{-24}\ \mathrm{A\cdot m^2} \tag{7.11}$$

est appelé *magnéton de Bohr.*

On sait que l'électron est doté d'un moment cinétique interne appelé *spin* $\mathcal{L}_s$ auquel est lié le moment magnétique de spin μ_s.

La projection de ce moment sur la direction de l'intensité du champ magnétique ne peut prendre que l'une de deux valeurs possibles: $\mu_{sH} = \pm\mu_B$.

Connaissant les moments magnétiques orbital et de spin des électrons distincts et en utilisant les règles de quantification spatiale, on peut déterminer le moment magnétique résultant M_J de l'atome tout entier. L'expression obtenue dans ce cas est analogue à (7.10):

$$M_J = g\mu_B \sqrt{J(J+1)}, \tag{7.12}$$

où J est un *nombre quantique interne* qui se détermine par les nombres quantiques orbital et de spin des électrons; g est un coefficient dit *facteur de Landé*. Il est égal à 1 pour un magnétisme purement orbital et à 2 pour un magnétisme dû uniquement au spin. Suivant les lois de la mécanique quantique, la projection de $\mathbf{M}_J$ sur la direction de $\mathbf{H}$ a pour valeur

$$M_{JH} = -m_J g\mu_B, \tag{7.13}$$

où

$$m_J = -J,\ -(J-1),\ \ldots,\ 0,\ 1,\ 2,\ \ldots,\ J \tag{7.14}$$

représente le *nombre quantique magnétique* de l'atome. Il résulte de (7.14) que pour J donné m_J peut prendre au total $2J+1$ valeurs.

Un atome de moment magnétique $\mathbf{M}_J$ possède dans le champ magnétique $\mathbf{H}$ une énergie magnétique

$$U_m = -\mu_0 H M_J \cos\theta = -\mu_0 H M_{JH}, \tag{7.15}$$

où θ est l'angle fait entre $\mathbf{M}_J$ et $\mathbf{H}$ (fig. 7.4, *c*). L'énergie U_m passe par sa valeur minimale pour $\theta = 0$. C'est pourquoi, tous les atomes, en tant que dipôles magnétiques, tendent à s'orienter dans le sens du champ. Le moment magnétique résultant d'une substance vaut la somme des projections des moments magnétiques des atomes distincts sur la direction de $\mathbf{H}$. Comme les valeurs de ces projections $M_{JH} = M_J \cos\theta = -m_J g\mu_B$, le problème du calcul quantitatif de l'aimantation de la substance se réduit au calcul de la valeur moyenne de M_{JH} correspondant à l'équilibre entre l'action orientatrice du champ et l'action désorientatrice de l'agitation thermique. Ce problème a été résolu par Langevin à l'aide des méthodes de la statistique classique et par Brillouin qui a utilisé les méthodes de la statistique quantique.

Considérons l'idée de la théorie quantique du paramagnétisme introduite par Brillouin. Comme il a été indiqué plus haut, le moment magnétique de l'atome peut être orienté dans le champ magnétique par $2J+1$ procédés auxquels correspondent $2J+1$ projections différentes du moment sur la direction du champ. La probabilité de réalisation de chacune de ces orientations est déterminée par la loi de Boltzmann $w = C \exp[\mu_0 M_{JH} H/(kT)]$, où C est un coeffi-

cient de proportionnalité. La valeur moyenne de M_{JH} est égale à

$$\langle M_{JH}\rangle = \frac{\sum_{-J}^{J} M_{JH} \exp\left[M_{JH}\mu_0 H/(kT)\right]}{\sum_{-J}^{J} \exp\left[M_{JH}\mu_0 H/(kT)\right]}. \tag{7.16}$$

Le calcul de ces sommes donne le résultat suivant:

$$\langle M_{JH}\rangle = gJ\mu_B B_J(\beta), \tag{7.17}$$

où

$$\beta = Jg\mu_B\mu_0 H/(kT), \tag{7.18}$$

$$B_J(\beta) = \frac{2J+1}{2J}\coth\frac{2J+1}{2J}\beta - \frac{1}{2J}\coth\frac{\beta}{2J}. \tag{7.19}$$

La fonction $B_J(\beta)$ est appelée *fonction de Brillouin.*

L'aimantation et la susceptibilité magnétique sont respectivement égales à

$$J_m = \langle M_{JH}\rangle n = ngJ\mu_B B_J(\beta), \tag{7.20}$$

$$\varkappa = \frac{ngJ\mu_B}{H} B_J(\beta), \tag{7.21}$$

où n est le nombre d'atomes par unité de volume. Pour $\beta \ll 1$,

$$B_J(\beta) \approx \frac{J+1}{3J}\beta$$

et

$$J_m = \frac{ng^2\mu_B^2 J(J+1)}{3kT}\mu_0 H, \tag{7.22}$$

$$\varkappa = \frac{ng^2\mu_B^2 J(J+1)\mu_0}{3kT}. \tag{7.23}$$

Il se trouve qu'en parfait accord avec l'expérience, J_m est proportionnel à H et inversement proportionnel à T. La formule (7.23) exprime la loi de Curie: $\varkappa = C/T$. La constante de Curie $C = ng^2\mu_B^2 J(J+1)\mu_0/(3k)$. Pour $\mu_0 H \approx 1$ T on a $\mu_B\mu_0 H \approx 10^{-23}$ J ; à $T = 300$ K on a $kT \approx 4\cdot 10^{-21}$ J. Ainsi la condition $\beta \ll 1$ est satisfaite pratiquement presque toujours. C'est seulement dans des champs très intenses et aux basses températures que $\beta \gg 1$ et la proportionnalité directe entre J_m et H est troublée; quand $\beta \to \infty$, $\coth\frac{2J+1}{2J}\beta \to 1$, $\coth\frac{\beta}{2J} \to 1$, $B_J(\beta) \to 1$ et l'aimantation atteint sa valeur de saturation:

$$J_m = J_s = ngJ\mu_B. \tag{7.24}$$

Paramagnétisme du gaz électronique. Le spin des électrons libres des métaux peut être orienté dans le champ magnétique soit suivant

le champ, soit dans le sens opposé. Ceci étant, chaque niveau de Landau (v. § 4.10) se décompose dans le champ magnétique en deux sous-niveaux. Comme l'énergie des électrons de spin orienté le long de l'intensité de champ est inférieure à l'énergie des électrons dont le spin est orienté dans le sens opposé à l'intensité de champ, le nombre d'électrons dans l'état d'équilibre thermique est plus grand au niveau inférieur qu'au niveau supérieur. C'est pour cette raison que le gaz électronique, pris dans son ensemble, acquiert un moment magnétique dirigé le long du champ et déterminant son paramagnétisme. Comme le montrent le calcul et l'expérience, le paramagnétisme du gaz électronique dans les métaux ne dépend pas de la température. Il détermine les propriétés paramagnétiques de nombreux métaux (Li, Na, Rb, K et autres).

Résonance paramagnétique. Si un paramagnétique placé dans un champ magnétique permanent est soumis à l'action d'un champ électromagnétique alternatif, on peut provoquer le passage des atomes des niveaux inférieurs aux niveaux plus élevés. En vertu des règles de sélection spectroscopiques, les seules transitions permises sont ici celles pour lesquelles le nombre quantique magnétique varie d'une unité : $\Delta m_J = \pm 1$, c'est-à-dire des transitions entre les sous-niveaux voisins présentant une différence d'énergie égale à $g\mu_B\mu_0 H_0$. De telles transitions peuvent être produites par un champ électromagnétique dont l'énergie des quanta

$$\hbar\omega = g\mu_B\mu_0 H_0. \tag{7.25}$$

C'est pourquoi lorsque la condition (7.25) est accomplie, on constate une absorption intense de l'énergie du champ électromagnétique.

La relation (7.25) exprime la condition de *résonance magnétique électronique* (RME). Le phénomène de résonance consiste en une absorption intense de l'énergie électromagnétique de fréquence déterminée par la substance. Cela explique le nom d'absorption *par résonance* qu'on donne à ce phénomène. Comme le montre l'expression (7.25) sa fréquence dépend de l'intensité du champ magnétique permanent H_0. Pour $\mu_0 H_0 = 1$ T et $g = 1$, la fréquence $\nu_{res} = 15$ GHz, ce qui correspond à une longueur d'onde $\lambda = 0{,}016$ m.

La condition de résonance (7.25) est relative à des atomes isolés, distincts, comportant des électrons non réunis en paires et possédant des moments magnétiques. Mais elle est également valable pour des corps constitués d'un grand nombre d'atomes si l'interaction magnétique entre eux est négligeable. Ces corps sont des paramagnétiques. C'est pourquoi la résonance des électrons non réunis en paires est appelée *résonance paramagnétique électronique* (RPE). Les premières expériences heureuses sur cette résonance ont été effectuées par Zavoïski en 1944. La fig. 7.5, *b* montre le schéma de principe de l'installation pour l'observation de ce phénomène. L'échantillon est placé dans le champ magnétique permanent $\mathbf{H}_0$ produit par les pièces

polaires N et S d'un électro-aimant. On soumet l'échantillon à l'action d'un champ électromagnétique à haute fréquence, produit par un générateur, de telle sorte que son vecteur intensité soit perpendiculaire à $\mathbf{H}_0$. En faisant varier la fréquence ω du champ électromagnétique ou l'intensité du champ permanent $\mathbf{H}_0$, on peut obtenir l'absorption par résonance de l'énergie à haute fréquence par l'échantillon.

On peut observer d'une manière analogue une *résonance magnétique nucléaire* (RMN) due au moment magnétique des noyaux. Par exemple, dans le cas des protons pour $\mu_0 H_0 = 1$ T la résonance nucléaire

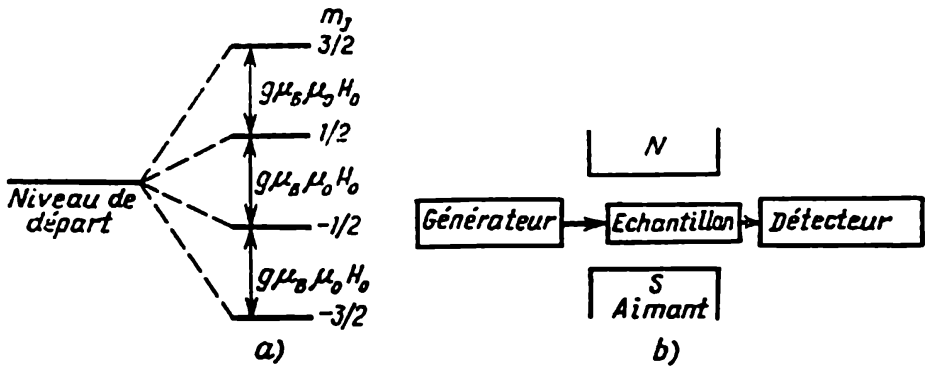

Fig. 7.5. Résonance magnétique:
a — décomposition du niveau d'énergie de départ d'un atome libre en sous-niveaux dans un champ magnétique; *b* — schéma de principe de l'installation pour observation d'une résonance paramagnétique.

s'observe à la fréquence $\nu_{res} = 8$ MHz, ce qui correspond à une longueur d'onde du rayonnement électromagnétique $\lambda \approx 10$ m. Les premières expériences sur la résonance nucléaire ont été effectuées par Torri et Pound (1945).

La résonance magnétique a reçu un nombre très grand d'applications pratiques. La résonance magnétique nucléaire constitue la méthode principale et la plus précise de mesure des moments magnétiques des noyaux atomiques. Elle a permis d'obtenir de nombreux renseignements sur la structure des liquides, des cristaux diélectriques, des métaux, des semiconducteurs et des polymères. Elle constitue aussi la méthode principale d'étude de la structure des molécules organiques. A l'aide de la résonance paramagnétique électronique on a étudié pour la première fois des états d'inversion de population des niveaux utilisés en électronique quantique, etc. Cette méthode permet d'étudier des particules comprenant des électrons non réunis en paires, ainsi que les processus auxquels elles participent. Ces particules sont représentées par les électrons de conduction, les radicaux libres, de nombreux atomes et ions. La méthode de la résonance paramagnétique électronique est employée avec succès pour l'étude du mécanisme des réactions chimiques, pour l'étude

de l'influence du rayonnement ionisant sur la substance et les tissus vivants, pour l'étude de l'état électronique des corps solides et dans de nombreuses autres branches de la technique et de la science.

§ 7.4. Nature du ferromagnétisme

Nature de spin du ferromagnétisme. Le ferromagnétisme n'est inhérent qu'à l'état solide de la substance. Les atomes libres des ferromagnétiques ne possèdent pas de propriétés magnétiques particulières quelconques. Les ordres de grandeur des moments magnétiques que possèdent les atomes d'éléments ferromagnétiques sont les mêmes que ceux des moments magnétiques des atomes d'éléments paramagnétiques. Mais l'induction dans les ferromagnétiques atteint, même dans des champs magnétiques faibles, des valeurs se chiffrant par quelques teslas alors que dans les paramagnétiques elle ne dépasse pas, dans les champs de même intensité, 10^{-5} T.

Pour expliquer les propriétés ferromagnétiques des ferromagnétiques, le physicien russe Rosing et le physicien français Weiss ont émis l'idée que dans les ferromagnétiques il existe un champ moléculaire interne sous l'effet duquel ils s'aimantent à saturation même en l'absence de champ magnétisant extérieur. Une telle *aimantation spontanée* ne se manifeste pas extérieurement parce que le corps se divise en domaines microscopiques distincts dont les moments magnétiques des atomes sont disposés dans chaque domaine parallèlement l'un à l'autre. Quant aux domaines eux-mêmes, ils sont orientés, l'un par rapport à l'autre, d'une façon suffisamment désordonnée de sorte que le moment magnétique résultant du ferromagnétique est nul. De telles régions d'aimantation spontanée sont appelées *domaines ferromagnétiques*. Actuellement il existe plusieurs méthodes expérimentales permettant une observation directe des domaines et la détermination de la direction de leur aimantation.

L'introduction du champ intramoléculaire a permis d'expliquer de nombreux événements observés dans les ferromagnétiques mais la nature du champ lui-même et la question de savoir quels moments magnétiques — orbitaux ou de spin — sont responsables du ferromagnétisme, sont restées pendant longtemps sans réponse. C'est seulement après les expériences effectuées avec le plus grand soin par Einstein et De Haas, Barnett, Ioffé et Kapitza qu'il a été démontré que le ferromagnétisme est dû aux moments magnétiques de spin des atomes. Aussi, ce sont précisément les moments magnétiques de spin des électrons, non compensés dans les atomes, qui constituent les *porteurs élémentaires* de ferromagnétisme.

Cette conclusion s'accorde avec la structure électronique des atomes d'éléments chimiques présentant le ferromagnétisme. Comme les moments magnétiques des couches complètement remplies sont nuls et, dans le métal, les électrons périphériques de valence sont

mis en commun seuls les éléments *de transition* possédant des couches internes *incomplètes* peuvent présenter le ferromagnétisme. Ces éléments sont les métaux de transition du groupe de fer dont la couche $3d$ est incomplète et les terres rares dont la couche $4f$ est incomplète. D'un autre côté, vu que les moments magnétiques orbitaux de ces couches sont pratiquement « congelés » par suite d'une forte interaction avec les atomes voisins et leur contribution aux propriétés magnétiques de la substance est bien faible, le ferromagnétisme des éléments faisant partie de ces groupes ne peut avoir son origine que dans les moments magnétiques de spin des atomes qui atteignent des valeurs élevées.

Il restait encore une question à élucider : quelles sont les forces qui obligent les moments magnétiques de spin des électrons de s'arranger parallèlement l'un à l'autre à l'intérieur du domaine ? On supposait d'abord que ces forces étaient de nature purement magnétique, qu'elles apparaissaient par suite d'une interaction ordinaire des moments magnétiques des spins (interaction spin-spin). L'énergie d'une telle interaction est de l'ordre de $U_m \approx -\mu_0 \mu_B^2 / a^3$, où a est la distance entre les atomes dans le réseau cristallin du ferromagnétique. En portant dans cette expression $\mu_B = 9{,}27 \cdot 10^{-24}$ A·m² et $a = 3 \cdot 10^{-10}$ m, on obtient $U_m \approx 10^{-23}$ J. Cela est inférieur de deux ordres de grandeurs environ à l'énergie d'agitation thermique des atomes du réseau à la température ordinaire qui tend à troubler la disposition ordonnée des spins. Cela montre que l'interaction magnétique des spins n'est pas en mesure d'assurer leur orientation parallèle caractéristique des ferromagnétiques à une température inférieure au point de Curie et que le champ moléculaire qui conduit à une telle orientation des spins doit être de nature *non magnétique*. Plus tard cette conclusion a été prouvée par des expériences directes de Dorfmann.

En 1928 Frenkel a émis l'idée que le champ moléculaire était de nature *électrostatique* et apparaissait par suite de l'*interaction d'échange* des électrons situés dans les couches internes incomplètes des atomes. Expliquons ceci sur l'exemple de liaison covalente dans la molécule d'hydrogène. Si, dans l'équation de Schrödinger, l'on ne tient pas compte des interactions magnétiques, ses solutions, c'est-à-dire les fonctions d'onde stationnaires, peuvent être représentées par un produit des fonctions $\psi(\mathbf{r}_1, \mathbf{r}_2)$, où $\mathbf{r}_1$ et $\mathbf{r}_2$ sont les coordonnées des électrons, par des fonctions dépendant uniquement des spins des électrons. Suivant le principe de Pauli, dans sa formulation mathématique, la fonction d'onde complète Ψ doit changer de signe en cas de permutation des électrons. Ce renversement de signe peut se produire par suite du changement de signe soit de la partie de spin de la fonction Ψ, soit de sa partie des coordonnées. Ainsi, les fonctions d'onde stationnaires des coordonnées pour la molécule d'hydrogène peuvent être symétriques et alors les spins

des électrons sont dirigés dans des sens opposés, mais elles peuvent aussi être antisymétriques et alors les spins des électrons sont orientés dans un même sens. Les fonctions d'onde des coordonnées différentes correspondent à des états d'énergie différente (électrostatique parce que les interactions magnétiques ont été complètement négligées). La différence d'énergie entre les états les plus bas caractérisés par des fonctions d'onde des coordonnées antisymétriques et symétriques détermine l'avantage énergétique de l'orientation parallèle des spins et donc aussi des moments magnétiques des électrons par rapport à l'orientation antiparallèle. Dans la molécule d'hydrogène cette différence est négative et les spins sont antiparallèles. Mais il se trouve que pour certains atomes cette différence est positive et dans ce cas il est énergétiquement plus avantageux que les moments de spin soient parallèles. C'est précisément ce cas qui est réalisé dans les ferromagnétiques.

Dans la théorie approchée, la variation de l'énergie d'un système électronique constitué de deux atomes ou de deux ions s'écrit en cas de changement mutuel de l'orientation de leurs moments de spin sous la forme suivante:

$$U_{éch} = -2J\,(\mathbf{S}_i \cdot \mathbf{S}_j). \tag{7.26}$$

La grandeur $U_{éch}$ est appelée *énergie d'échange*; J est l'intégrale d'échange; $\mathbf{S}_i$, $\mathbf{S}_j$ sont les moments de spin résultants des atomes interagissants. L'intégrale d'échange peut être aussi bien positive que négative. Son signe dépend du rapport entre la dimension d des couches électroniques participant à la formation de la liaison d'échange et la distance a entre les atomes (paramètre du réseau). La figure 7.6, *a* montre à titre d'exemple la courbe de variation de J en fonction du rapport a/d pour quelques éléments. Comme il est visible sur cette figure, pour des éléments dont le rapport $a/d > 1{,}5$ (Fe, Co, Ni, Gd) l'intégrale d'échange est positive et pour des éléments dont $a/d < 1{,}5$ (Mn, Cr) elle est négative et son module est d'autant plus grand que le rapport a/d est plus petit.

Comme il a déjà été dit, le signe de l'intégrale d'échange détermine celle des orientations des spins des électrons participant à la formation de la liaison d'échange — parallèle ou antiparallèle — qui est avantageuse. Lorsque le signe de l'intégrale d'échange est positif ($J > 0$), l'énergie d'échange est négative et l'énergie du système pris dans son ensemble diminue dans le cas où les spins des atomes $\mathbf{S}_i$ et $\mathbf{S}_j$ sont parallèles l'un à l'autre ($\mathbf{S}_i \downarrow\downarrow \mathbf{S}_j$). Comme le montre la figure 7.6, *a*, cette situation doit se présenter pour le fer, le cobalt et le nickel. Pour les spins des électrons situés sur les couches $3d$ incomplètes des atomes il est énergétiquement avantageux de se disposer parallèlement l'un à l'autre, en assurant une saturation magnétique limite des domaines (fig. 7.6, *b*). C'est pourquoi les Fe, Co et Ni sont des ferromagnétiques. Pour les Mn et Cr, $J < 0$,

de sorte que c'est la disposition antiparallèle des spins des atomes voisins du réseau qui est avantageuse. Cela explique pourquoi les Mn et Cr ne possèdent pas de propriétés ferromagnétiques. Toutefois si l'on augmente la constante du réseau de Mn de façon que le rapport a/d devienne de l'ordre de 1,5, on peut s'attendre à ce que le manganèse devienne ferromagnétique. L'expérience le confirme. C'est ainsi que l'introduction dans le Mn d'une petite quantité d'azote a pour effet d'augmenter le paramètre du réseau et de faire apparaître le ferromagnétisme. Sont aussi ferromagnétiques les alliages Mn-Cu-Al

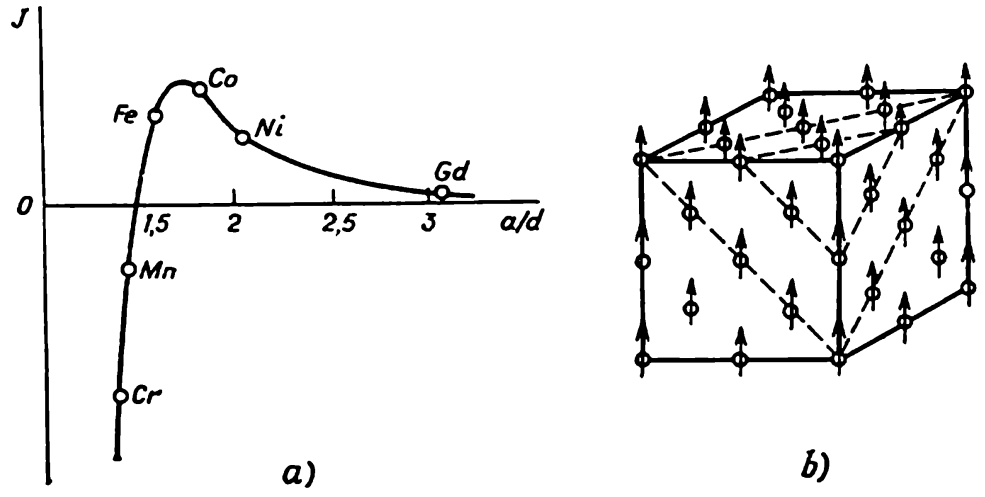

Fig. 7.6. Aimantation spontanée d'un ferromagnétique:
a — variation de l'intégrale d'échange en fonction du rapport du paramètre *a* du réseau au diamètre *d* des couches internes incomplètes de certains éléments; *b* — orientation des spins des électrons des couches électroniques internes incomplètes dans les ferromagnétiques.

(*alliages de Geissler*) et les composés MnSb, MnBi et autres dans lesquels les atomes de manganèse sont séparés par des distances plus grandes que dans le réseau cristallin du manganèse pur.

Ainsi, *la présence dans l'atome de couches internes incomplètes et le signe positif de l'intégrale d'échange déterminant une orientation parallèle des spins sont des conditions dont la réalisation fait apparaître le ferromagnétisme.*

Si les spins des atomes en interaction sont égaux en module ($S_i = S_j = S$) et forment entre eux un angle θ, l'expression (7.26) peut se récrire sous la forme

$$U_{éch} = -2JS^2 \cos \theta. \qquad (7.27)$$

Anisotropie cristallographique magnétique. Suivant la théorie simplifiée développée plus haut pour l'énergie de l'état des électrons n'est important que le parallélisme ou l'antiparallélisme des moments de spin des atomes voisins dans le cristal qui détermine respectivement l'antisymétrie ou la symétrie de la partie dépendant des coor-

données de leur fonction d'onde. Or, en réalité, comme le montrent l'expérience et la théorie tenant compte de l'interaction spin-orbite, il n'est pas du tout indifférent comment sont dirigés les moments de spin parallèles ou antiparallèles par rapport aux axes cristallographiques. Dans le cristal il existe des directions d'orientation des moments magnétiques d'énergie d'échange minimale dites directions *de facile aimantation* et des directions d'énergie maximale dites directions *de difficile aimantation*. Par exemple, dans le cristal de fer la direction de facile aimantation est la direction ⟨100⟩ et celle de difficile aimantation la direction ⟨111⟩.

La variation de l'énergie d'un cristal par suite d'une rotation du vecteur aimantation par rapport à la direction de facile aimantation est appelée *énergie d'anisotropie magnétique*. Pour les cristaux cubiques cette énergie peut s'écrire sous la forme

$$U_c = K_1 (\alpha_1^2\alpha_2^2 + \alpha_2^2\alpha_3^2 + \alpha_3^2\alpha_1^2) + K_2\alpha_1^2\alpha_2^2\alpha_3^2, \qquad (7.28)$$

où K_1 et K_2 sont des *constantes d'anisotropie* et α_1, α_2, α_3 sont les cosinus directeurs du vecteur aimantation par rapport aux axes du cristal. Les constantes d'anisotropie dépendent de la température et tombent à zéro quand $T \to \Theta_C$. Pour le fer à la température ordinaire $K_1 = 4{,}2 \cdot 10^4$ J/m³, $K_2 = 1{,}5 \cdot 10^4$ J/m³. Pour les cristaux hexagonaux

$$U_c = K_1' \sin^2\theta + K_2' \sin^4\theta, \qquad (7.29)$$

où θ est l'angle formé entre l'aimantation spontanée J_m et l'axe hexagonal du cristal qui coïncide avec la direction de facile aimantation. Pour le cobalt, à la température ordinaire, $K_1' = 4{,}1 \times \times 10^5$ J/m³, $K_2' = 1{,}0 \cdot 10^5$ J/m³.

Antiferromagnétisme. Comme il a déjà été indiqué, lorsque le signe de l'intégrale d'échange est négatif, c'est l'orientation antiparallèle des spins des atomes voisins du réseau qui devient énergétiquement avantageuse. Dans ce cas lui aussi la disposition des spins peut être ordonnée mais l'aimantation spontanée est nulle parce que les moments magnétiques de spin des atomes voisins du réseau sont orientés deux à deux en sens contraire et se compensent. La figure 7.7, *a* montre à titre d'exemple la structure magnétique de MnO déterminée par les méthodes de spectroscopie neutronique (on n'a représenté sur la figure que les atomes de Mn magnétoactifs). Elle peut être considérée comme une structure complexe, composée de deux sous-réseaux aimantés dans des sens contraires. Une telle structure ne peut exister qu'à des températures inférieures à une certaine température qu'on appelle *température (point) de Curie antiferromagnétique* ou *température (point) de Néel* T_N.

A 0 K, les moments magnétiques des sous-réseaux se neutralisent mutuellement de sorte que le moment magnétique résultant d'un antiferromagnétique placé dans un champ extérieur dirigé le long

des moments magnétiques des sous-réseaux est nul. Lorsque la température s'élève, l'orientation antiparallèle des spins se détruit progressivement et l'aimantation de l'antiferromagnétique augmente

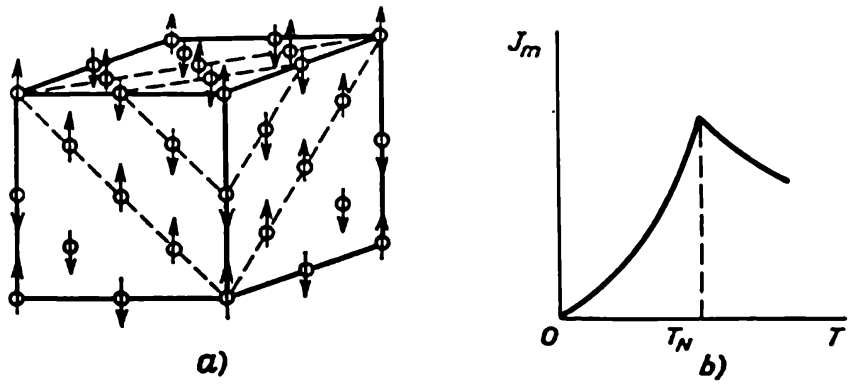

Fig. 7.7. Structure magnétique d'un antiferromagnétique (MnO) (*a*) [et variation de son aimantation en fonction de la température (*b*).

(fig. 7.7, *b*) pour atteindre sa valeur maximale au point de Néel où la disposition ordonnée des spins est totalement anéantie et l'antiferromagnétique se transforme en paramagnétique. Si la température augmente encore, l'aimantation diminue comme pour tout paramagnétique.

Notions sur le ferrimagnétisme. Dans les antiferromagnétiques les moments magnétiques des sous-réseaux sont égaux en module et de signes contraires de sorte qu'ils se neutralisent complètement. Pourtant des cas peuvent se présenter où les moments magnétiques des sous-réseaux ne sont pas égaux, par exemple du fait que les atomes constituant ces sous-réseaux sont aux nombres différents ou de nature différente (fig. 7.8, *a*). Dans ce cas il apparaît une différence non nulle entre les moments magnétiques des sous-réseaux qui provoque une aimantation spontanée du cristal. Un tel antiferromagnétisme *non compensé* est appelé *ferrimagnétisme*.

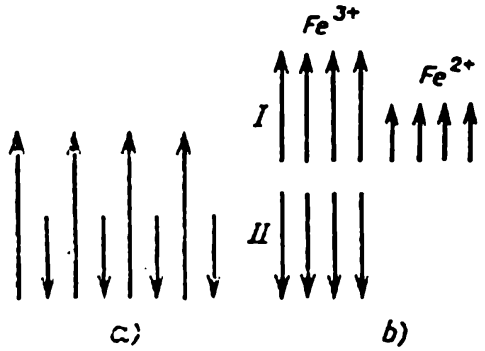

Fig. 7.8. Schéma de la disposition des moments magnétiques dans les ferrimagnétiques (*a*); disposition des moments magnétiques des ions Fe^{3+} et Fe^{2+} dans le réseau de la magnétite $FeO \cdot Fe_2O_3$: *I* — dans les nœuds octaédriques, *II* — dans les nœuds tétraédriques.

Un ferrimagnétique se comporte en apparence comme un ferromagnétique mais par suite de leur différence interne, la variation de leur aimantation spontanée avec la température peut être tout à fait différente. Comme un exemple de ferrimagnétique on peut indiquer

la magnétite $FeO \cdot Fe_2O_3$. Elle possède la structure cristalline de la spinelle dont la maille élémentaire cubique contient 8 ions Fe^{2+} et 16 Fe^{3+} qui occupent des nœuds tétraédriques (II) et octaédriques (I) correspondants du réseau. Les directions des moments magnétiques de ces ions sont montrées schématiquement sur la figure 7.8, *b*. Comme il est visible sur cette figure, les moments magnétiques des ions Fe^{3+} sont deux à deux antiparallèles l'un à l'autre de sorte qu'ils se neutralisent mutuellement. Aussi le moment magnétique résultant du réseau se détermine-t-il seulement par les moments magnétiques non compensés des ions Fe^{2+}. Les ions bivalents de fer peuvent être remplacés par des ions d'autres métaux bivalents, par exemple Mg, Ni, Co, Cu, etc., de sorte que la formule générale des substances de cette classe qui ont reçu le nom de *ferrites* est $MeO \cdot Fe_2O_3$, où Me désigne un métal bivalent.

Ce sont des ferrites dits mixtes, dont la composition comprend les oxydes de plusieurs métaux bivalents, par exemple NiO et ZnO, MnO et MgO, MgO et ZnO, etc., qui présentent un grand intérêt. En changeant la nature de ces oxydes et leurs rapports, on peut faire varier dans de larges limites les propriétés magnétiques et autres des ferrites. Notamment, on peut obtenir des ferrites magnétiquement doux, des ferrites magnétiquement durs, des ferrites à cycle d'hystérésis rectangulaire, etc.

Une particularité remarquable des ferrites est leur haute résistivité électrique qui est de 10^5 à 10^{15} fois plus grande que la résistivité des ferromagnétiques métalliques. Cette particularité a permis de vaincre une difficulté qui s'est présentée dans la technique des hautes et très hautes fréquences et qui paraissait insurmontable en matière d'utilisation des matériaux magnétiques. Le fait est que dans la majeure partie de dispositifs radiotechniques utilisant des champs magnétiques, les bobines parcourues par un courant comportent, pour l'amplification de ces champs, des noyaux (circuits magnétiques) faits en matériaux ferromagnétiques. Si les bobines sont destinées à être alimentées en courant continu, les noyaux peuvent être fabriqués en ferromagnétique continu, par exemple en fer, en permalloy, etc. Si les bobines doivent être alimentées par des courants alternatifs et surtout de fréquence élevée, les noyaux de ce type deviennent inutilisables parce que leur magnétisation alternative fait naître des courants de Foucault très forts, qui non seulement augmentent les pertes d'énergie et réduisent le rendement des dispositifs, mais peuvent échauffer le noyau à un tel point que le dispositif cesse de fonctionner ou peut même être mis hors d'état de service. Aussi, les noyaux sont-ils réalisés à partir des feuilles minces et de petites particules de ferromagnétiques isolées l'une de l'autre. Cela a permis de réduire considérablement les courants de Foucault mais n'a pas surmonté toutes les difficultés liées aux pertes, à l'effet de peau, etc., dont la manifestation est particulièrement sensible aux

hautes et très hautes fréquences. Le succès n'a été réalisé qu'avec l'apparition des ferrites dans lesquels les propriétés magnétiques des ferromagnétiques sont jointes aux propriétés électriques des diélectriques.

Pour la construction des noyaux pour les bobines, les transformateurs, les antennes à noyau magnétique et de divers circuits magnétiques à haute fréquence, on utilise des ferrites magnétiquement doux caractérisés par de hautes valeurs de la perméabilité magnétique μ, de l'induction de saturation B_s et de l'induction rémanente B_r mais par de petites valeurs de la force coercitive H_c et de la tangente de pertes tg δ. Ce sont des ferrites de manganèse-zinc, de nickel-zinc, de lithium-zinc et autres.

Les ferrites magnétiquement durs (de fer-cobalt, de fer-baryum et autres) possédant une grande force coercitive H_c et une grande induction rémanente B_r sont utilisés pour la fabrication des aimants permanents. La haute résistivité de tels ferrites permet de les utiliser dans la technique des hyperfréquences pour les systèmes de prémagnétisation.

§ 7.5. Structure en domaines des ferromagnétiques

Formation de domaines. Comme il a déjà été indiqué, un ferromagnétique à l'état non aimanté se divise spontanément en domaines aimantés à saturation grâce à l'orientation parallèle de leurs moments

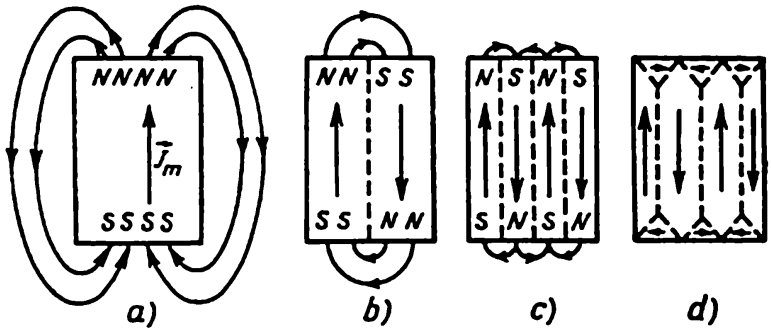

Fig. 7.9. Division d'un cristal ferromagnétique en domaines.

magnétiques de spin qui se produit sous l'action des forces d'échange. La cause d'une telle division est facile à comprendre si l'on examine la figure 7.9. La figure 7.9, *a* montre un cristal monodomaine spontanément aimanté dans la direction de $\mathbf{J}_m$. Un tel cristal possède une haute énergie magnétique due à la formation sur ses faces terminales de « pôles magnétiques ». Lorsque le cristal est divisé en deux domaines dont les vecteurs aimantation spontanée sont opposés l'un à l'autre (fig. 7.9, *b*), l'énergie magnétique du cristal devient près de deux fois plus faible parce qu'une partie du flux magnétique sortant d'un des domaines se ferme sur l'autre domaine ; si le cristal est divisé

en quatre domaines (fig. 7.9, *c*), l'énergie magnétique diminue d'environ quatre fois, etc.

Ainsi, la division d'un cristal ferromagnétique en domaines est une conséquence de la tendance du système à diminuer son énergie libre (magnétique). Toutefois, cette division ne peut pas se produire indéfiniment parce que l'apparition de frontières entre les domaines voisins *I* et *II* dont l'angle formé entre les spins augmente jusqu'à 180° (fig. 7.10, *a*) doit conduire inévitablement à une augmentation de l'énergie d'échange. La division se poursuit jusqu'à l'instant où la diminution de l'énergie magnétique provoquée par la division du

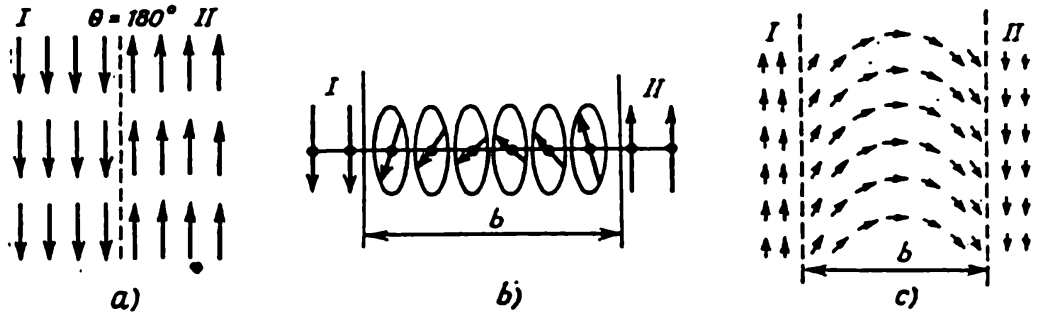

Fig. 7.10. Frontières entre les domaines (parois de domaines) :
a — lorsque les spins à la surface de séparation de deux domaines sont de sens contraires, l'énergie d'échange de la surface de séparation est maximale ; *b*, *c* — variation progressive de la direction du vecteur aimantation lors du passage d'un domaine à un autre et formation de parois des domaines de Bloch (*b*) et de Néel (*c*).

cristal en domaines est compensée par l'augmentation de l'énergie d'échange des frontières de séparation entre les domaines. Une division ultérieure n'est pas avantageuse énergétiquement et ceci détermine la limite inférieure de la dimension du domaine. Comme le montrent le calcul et l'expérience, pour le fer la dimension transversale du domaine est de ~0,01 à 0,1 μm.

Lors du passage par bond du domaine ayant une direction des spins à un domaine ayant une direction opposée des spins (fig. 7.10, *a*) l'énergie d'échange des atomes situés de part et d'autre de la surface de séparation des domaines atteint sa valeur maximale possible égale à $-2JS^2 \cos 180° = 2JS^2$. C'est pourquoi l'énergie d'échange de la frontière entre les domaines considérés comme un tout atteindra elle aussi sa valeur maximale. Comme le montrent le calcul et l'expérience, le passage énergétiquement plus avantageux est un passage progressif de l'orientation des spins dans un domane à l'orientation opposée dans un autre domaine qui se produit sur une certaine région étendue appelée *paroi de domaine*. Les structures de deux types de telles parois — *paroi de Bloch* et *paroi de Néel* — sont représentées sur les figures 7.10, *b* et *c*. Des parois de type plus complexe peuvent elles aussi exister. Plus épaisse est la paroi d'un domaine, plus faible est la valeur totale dont varie l'énergie d'échange lors du pas-

sage d'un domaine à un autre. Il semblerait que pour cette raison l'épaisseur de la paroi de domaine devrait croître indéfiniment. Mais l'anisotropie magnétique s'y oppose. Le cristal de fer comporte six directions de facile aimantation [100]. Lors de l'aimantation spontanée les moments magnétiques de spin se disposent suivant l'une de ces directions de sorte que l'énergie d'anisotropie magnétique a sa valeur minimale. Dans les parois de domaines les moments magnétiques de spin s'écartent de la direction de facile aimantation, l'énergie d'anisotropie magnétique augmente et cela d'autant plus fortement que la paroi est plus épaisse. Aussi, l'épaisseur des parois augmente-t-elle jusqu'à des valeurs pour lesquelles la diminution de l'énergie d'échange provoquée par leur apparition n'est pas compensée par la croissance de l'énergie d'anisotropie magnétique. Le calcul montre que pour le fer la paroi de Bloch doit avoir une épaisseur de l'ordre de 300 distances atomiques.

Une diminution de l'énergie magnétique, encore plus grande que celle de la figure 7.9, *c*, peut être obtenue si les domaines verticaux sont fermés par des domaines horizontaux cunéiformes (fig. 7.9, *d*). Pourtant, vu que dans la direction de facile aimantation le fer s'allonge légèrement (*magnétostriction*), les domaines horizontaux doivent être élastiquement comprimés. *La structure magnétique d'équilibre est atteinte lorsque la diminution de l'énergie magnétique par suite de l'apparition des domaines de fermeture est compensée par l'augmentation de l'énergie élastique du cristal provoquée par la déformation de ces domaines.*

La structure en domaines dépend fortement de la forme de l'échantillon. Si l'échantillon a par exemple la forme d'une tige étroite dont l'axe de facile aimantation est dirigé suivant sa longueur, il n'est pas avantageux pour lui de se diviser en domaines. La structure cristalline de l'échantillon a elle aussi une grande importance.

Aimantation des ferromagnétiques. Les processus d'aimantation des ferromagnétiques s'expliquent pour une large part par le changement de leur structure en domaines dans un champ extérieur. Considérons par exemple le processus d'aimantation d'un monocristal dans lequel les domaines sont orientés, en l'absence de champ extérieur, de telle sorte que le moment magnétique du ferromagnétique tout entier est nul (fig. 7.11, *a*). Lorsque le ferromagnétique est placé dans un champ magnétique **H**, il s'aimante en acquérant un moment magnétique non nul. Suivant le caractère des phénomènes physiques se déroulant dans le ferromagnétique le processus de magnétisation peut être subdivisé en trois stades.

1. *Processus de déplacement des frontières.* Plaçons le cristal représenté sur la figure 7.11, *a* dans un champ magnétique. L'orientation du vecteur $\mathbf{J}_m$ des différents domaines par rapport à **H** n'est pas la même : l'angle minimal avec **H** forme le vecteur $\mathbf{J}_m$ du premier domaine et l'angle maximal fait $\mathbf{J}_m$ du troisième domaine. Lorsque

l'intensité de H augmente c'est la croissance du domaine *1*, dont l'orientation est la plus favorable, qui est énergétiquement avantageuse au détriment des domaines *2*, *3* et *4* (fig. 7.11, *b*). Cette croissance se fait par déplacement des frontières des domaines. C'est pourquoi le premier stade de magnétisation a reçu le nom de *processus de déplacement*. Il se produit jusqu'à l'instant où le premier domaine

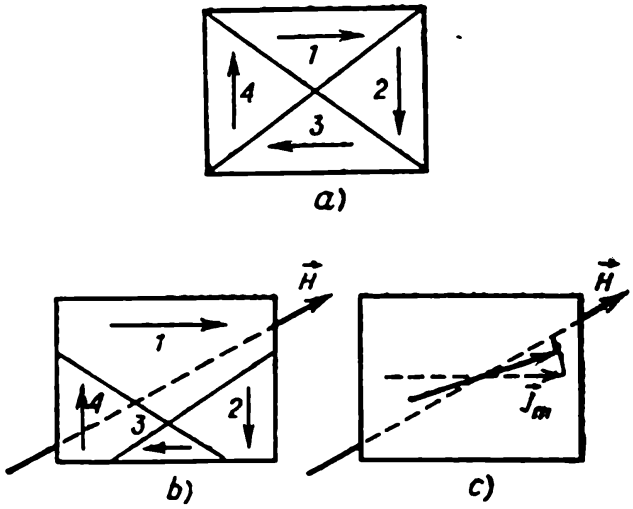

Fig. 7.11. Processus se déroulant lors de l'aimantation d'un cristal ferromagnétique :

a — surfaces de séparation de quatre domaines en lesquels s'est divisé le monocristal. Les flèches indiquent l'orientation du vecteur J_m ; *b* — déplacement des frontières et croissance du domaine le plus favorablement orienté *1* lorsque l'intensité du champ magnétisant H augmente; *c* — rotation du vecteur aimantation J_m dans le sens vers le champ H (la flèche en traits interrompus indique le vecteur J_m après la fin du processus de déplacement des frontières des domaines et la flèche en traits pleins représente le vecteur J_m tourné d'un certain angle par rapport au champ H).

s'étend sur tout le cristal. La figure 7.12 montre la courbe d'aimantation d'un monocristal. C'est le tronçon *OA* de cette courbe qui correspond au processus de déplacement. Dans des champs de faible intensité H l'aimantation se fait progressivement et de façon réversible tandis que dans des champs d'intensité moyenne elle se produit par bond et de façon irréversible, en déterminant un effet dit *effet Berkhausen*.

2. *Processus de rotation*. Si l'intensité du champ augmente encore, le vecteur aimantation spontanée $\mathbf{J}_m$ commence à tourner dans le sens vers le champ (fig. 7.11, *c*). Le processus d'aimantation se déroule ici beaucoup plus lentement qu'au premier stade et cesse lorsque le vecteur $\mathbf{J}_m$ se dispose le long du champ. A cet instant, l'aimantation atteint la valeur de saturation technique (tronçon *AB* sur la courbe de la figure 7.12).

3. *Paraprocessus*. Dès que l'état de saturation technique est atteint, la croissance de l'aimantation avec l'augmentation de H diminue brusquement sans cesser complètement. Cela s'explique par le fait qu'à une température différente de 0 K tous les spins ne sont

pas orientés de façon rigoureusement parallèle l'un à l'autre. Un champ extérieur intense diminue l'action de désorientation de l'agitation thermique et conduit à une augmentation ultérieure, bien que faible, de J_m.

Hystérésis magnétique. La non-coïncidence de la courbe d'aimantation répétée d'un ferromagnétique avec la courbe *OA* (fig. 7.3) de première aimantation, qui conduit à l'hystérésis, est due à des processus irréversibles qui se déroulent lors de l'aimantation. Examinons brièvement l'essence de ces processus.

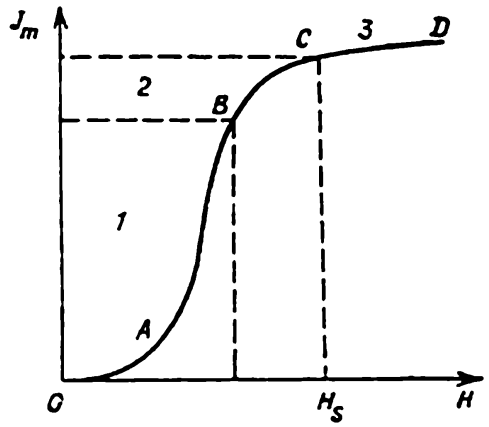

Fig. 7.12. Courbe d'aimantation des ferromagnétiques:

1 — région correspondant au déplacement des frontières des domaines; *OA* — déplacement réversible, *AB* — déplacement irréversible; *2* — région correspondant à la rotation du vecteur aimantation; *3* — région correspondant au paraprocessus.

1. *Irréversibilité lors du déplacement des frontières des domaines.* La présence dans un ferromagnétique de différentes hétérogénéités — des impuretés, des inclusions non magnétiques, des régions de contraintes, etc., peut avoir une forte influence sur l'énergie des parois de Bloch, en l'augmentant ou en la diminuant, c'est-à-dire en créant pour ces parois des puits de potentiel qu'elles traversent lors de leur déplacement au premier stade de l'aimantation. Lors de la désaimantation une partie de parois peut rester dans ces puits de sorte que les domaines qui étaient aimantés le long du champ sont conservés aussi après la suppression du champ, en assurant une aimantation rémanente B_r (v. fig. 7.3). Pour compenser cette aimantation, il est nécessaire d'appliquer un champ H_c de sens opposé. En réglant les facteurs qui déterminent la courbe d'aimantation et de désaimantation, on peut faire varier dans de larges limites la forme et les dimensions du cycle d'hystérésis. Pour des ferromagnétiques homogènes ne présentant qu'un nombre minimal de défauts, le cycle d'hystérésis peut être très étroit.

2. *Retard de croissance des germes de renversement de l'aimantation.* Lorsque la saturation technique est atteinte, l'aimantation des domaines coïncide pratiquement avec la direction du champ magnétisant. Si le sens du champ est inversé, pour que l'aimantation du ferromagnétique commence à se renverser, il doit y prendre naissance des *germes de renversement de l'aimantation* dont la croissance détermine le renversement de l'aimantation du ferromagnétique dans son ensemble. L'apparition de tels germes exige l'existence d'un *champ* suffisamment intense dit *de départ* $H_{dép}$. C'est pourquoi si, la saturation technique étant atteinte, on diminue jusqu'à zéro le champ

magnétisant et on applique ensuite un champ progressivement croissant de sens opposé, le renversement d'aimantation ne se produira pas et l'aimantation rémanente restera inchangée jusqu'à l'instant où le champ d'intensité inverse atteint la valeur de départ $H_{dép}$. Pour cette valeur de l'intensité il se crée dans l'échantillon des germes de renversement d'aimantation qui croissent à une grande vitesse (centaines de m/s), le ferromagnétique change la direction de l'aimantation de façon presque instantanée et le cycle d'hystérésis a une forme pratiquement rectangulaire. Notamment, les échantillons de permalloy uniaxialement contraints ont un tel cycle d'hystérésis. Mais dans le cas général, étant donné l'hétérogénéité de l'échantillon et la présence de champ démagnétisant, les germes de renversement de l'aimantation peuvent apparaître dans différents endroits du ferromagnétique et dans des champs extérieurs les plus variés. Aussi le cycle d'hystérésis peut-il différer fortement de la forme rectangulaire.

3. *Aimantation des particules hautement dispersives.* Lorsque les dimensions des particules ferromagnétiques diminuent, la part de l'énergie des parois de Bloch augmente. Aussi, à partir d'une certaine dimension des particules, devient-il énergétiquement avantageux que ces particules soient dans l'état monodomaine. Pour des particules sphériques de fer, le passage à la structure magnétique monodomaine s'effectue pour un rayon des particules égal à $\sim 0{,}1$ μm. Comme le domaine est aimanté à saturation, chacune de telles particules représente un petit aimant permanent. Si on répartit un grand nombre de telles particules dans un milieu non ferromagnétique et qu'on les aimante dans une direction, on peut obtenir un aimant permanent de forte intensité. A cet effet, on utilise généralement des alliages à plusieurs composants qui dégagent, lorsque la température baisse, des phases hautement dispersives dont une au moins est magnétique. Comme exemple de tels alliages, on peut indiquer les alliages à plusieurs composants de Al, Ni, Fe avec des additifs de silicium et de cobalt (alliages alni, alnisi, alnico et autres). Les particules magnétiques de ces alliages ont le plus souvent la forme de minuscules bâtonnets dont les dimensions sont si petites qu'ils ont la structure monodomaine avec l'axe de facile aimantation dirigé le long de l'axe du bâtonnet. Si l'alliage est refroidi dans un champ magnétique, les domaines-bâtonnets s'orientent par leurs axes le long du champ en produisant ainsi une anisotropie de structure et d'aimantation. Une telle structure possède un cycle d'hystérésis rectangulaire qui se caractérise par une forte aimantation rémanente et une haute force coercitive, ce qui est précisément important pour la réalisation des aimants permanents.

§ 7.6. Applications pratiques du ferromagnétisme

Mémoires magnétiques à ferrites. Actuellement le ferromagnétisme est d'application courante dans les appareils de calcul électronique. Les ferrites à cycle d'hystérésis rectangulaire sont utilisés pour la fabrication des mémoires magnétiques des ordinateurs, des amplificateurs magnétiques, des lignes à retard, etc. Un des paramètres principaux caractérisant l'applicabilité des ferrites à ces fins est le *facteur de rectangularité* du cycle d'hystérésis, égal au rapport de l'induction rémanente B_r à l'induction maximale B_{max} mesurée pour $H = 5H_c$:

$$K_{rect} = B_r/B_{max}.$$

Il faut que K_{rect} soit aussi voisin que possible de l'unité. En outre, le ferrite doit avoir une faible force coercitive. Les ferrites de magnésium-manganèse ayant $K_{rect} \approx 0{,}88$ à $0{,}96$ et $H_c = 8$ à 40 A/m présentent en particulier de telles propriétés.

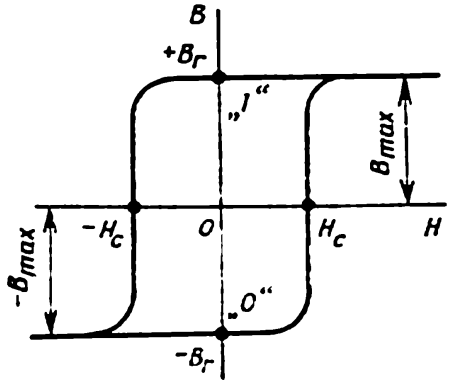

Fig. 7.13. Cycle d'hystérésis rectangulaire.

Le principe physique sur lequel est fondée la «mémoire» magnétique est le suivant. Supposons qu'un ferrite à cycle d'hystérésis rectangulaire est aimanté jusqu'à B_{max} par un champ H dirigé de gauche à droite (fig. 7.13). Lorsque ce champ diminue jusqu'à zéro, l'induction diminue jusqu'à la valeur de B_r qui, dans un cycle d'hystérésis rectangulaire, diffère peu de B_{max}. Lorsque le sens du champ H est inversé, l'induction reste presque inchangée jusqu'à $H = -H_c$. Pour $H = -H_c$ l'induction change de signe, par saut, en atteignant une valeur presque limite $-B_{max}$ qui varie peu lors de la croissance ultérieure de H. Si on diminue maintenant ce champ, l'induction rémanente du ferrite sera, pour $H = 0$, égale à $-B_r$. Ainsi, pour une intensité du champ extérieur $H = 0$ le ferrite peut se trouver dans deux états stables: avec $B = +B_r$ et $B = -B_r$ suivant son aimantation antérieure. C'est sur cette propriété des ferrites de «mémoriser» l'état d'aimantation antérieur qu'est fondé le fonctionnement des mémoires magnétiques.

La figure 7.14, *a* montre le schéma de principe d'une cellule de mémoire à noyau en ferrite. Le noyau porte trois enroulements: d'écriture *1*, d'interrogation *2* et de lecture *3*. L'écriture se fait dans le système binaire: « 1 » et « 0 ». Pour l'écriture de « 1 » on envoie dans l'enroulement *1* une impulsion de courant capable d'aimanter le noyau dans le sens positif jusqu'à B_{max}. Quand l'action de l'impulsion cesse, l'induction diminue jusqu'à $+B_r$ (v. fig. 7.13). Pour

l'écriture de « 0 » on fait passer dans l'enroulement *1* une impulsion de courant de sens opposé après l'action de laquelle l'induction rémanente du noyau devient égale à $-B_r$. Pour la lecture de l'information écrite, on fait passer dans l'enroulement *2* une impulsion de courant qui produit un champ magnétique $H_{int} > H_c$ dans le sens qui fait passer le noyau dans l'état « 0 ». Si l'information inscrite dans la cellule correspondait à « 1 », l'impulsion d'interrogation, qui change l'aimantation du noyau de l'état $+B_r$ en état $-B_r$, y produit un flux variable d'induction magnétique qui induit une

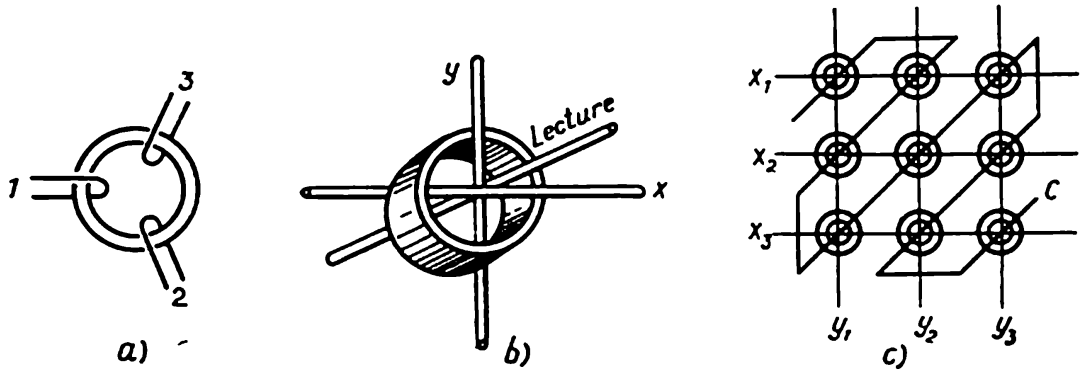

Fig. 7.14. Principe de construction des cellules de mémoire (*a*, *b*) et matrices de mémoire à noyaux en ferrite (*c*).

impulsion de courant dans l'enroulement de lecture *3*. Plus grande est la distance entre les points $+B_r$ et $-B_r$, plus brusque est le passage de $+B_r$ à $-B_r$, c'est-à-dire plus manifeste est la rectangularité du cycle d'hystérésis, plus intense est l'impulsion de courant induite dans l'enroulement *3* par suite du renversement d'aimantation. C'est pourquoi on n'emploie pour la réalisation des mémoires que des ferrites à cycle d'hystérésis rectangulaire.

Si c'est un « 0 » qui était enregistré dans la mémoire, le renversement d'aimantation du noyau par le courant d'interrogation ne se produit pas et aucun signal n'est induit dans l'enroulement de lecture. Chaque cellule de mémoire contient un élément d'information (un bit) signifiant l'une de deux affirmations possibles : « oui » (« 1 ») ou « non » (« 0 »).

La figure 7.14, *b* montre le schéma d'une cellule de mémoire dont la réalisation est un peu différente de celle de la cellule représentée par la figure 7.14, *a*. Les enroulements *1*, *2*, *3* sont remplacés ici par trois fils (barres) : x, y et une barre de lecture, qui passent par un anneau en ferrite. Pour l'écriture de l'information on fait passer par les fils x et y des impulsions de courant $I_x = I_0/2$ et $I_y = I_0/2$, où I_0 est l'intensité du courant nécessaire pour renverser l'aimantation du noyau. Aussi le renversement de l'aimantation du noyau ne peut-il être produit que par l'action simultanée des impulsions I_x et I_y de même sens.

A partir des cellules de mémoire on compose des matrices (l'une d'elles est schématisée sur la figure 7.14, *c*). La barre de lecture L est commune à toutes les cellules de la matrice. Pour l'écriture de « 1 » par exemple dans la cellule supérieure de droite, on fait passer dans x_1 et y_3 des impulsions de courant $I_0/2$ de même sens dans chaque barre. Pour l'écriture de « 0 » ou pour la lecture de « 1 » on fait passer par ces barres des impulsions de même sens $-I_0/2$ qui renversent l'aimantation du noyau et induisent un signal électrique dans la barre de lecture. En procédant de cette façon, on peut ranger l'information dans toute cellule de mémoire et ensuite l'utiliser.

La dimension des noyaux en ferrite utilisés dans les mémoires modernes des calculateurs électroniques est réduite à 0,3 mm pour le diamètre extérieur et à 0,15 mm pour le diamètre intérieur.

Films magnétiques minces. Par films magnétiques minces on entend des couches de substance magnétique épaisses de $\sim$ 0,1 μm portées sur un support non magnétique. Ce sont les films de permalloy contenant $\sim$ 20 % Ni et $\sim$ 80 % Fe qui ont reçu le plus d'utilisations. Ils sont déposés sur un support par les méthodes d'évaporation sous vide, de pulvérisation cathodique et de pulvérisation par faisceau ionique.

Dans des films d'épaisseur inférieure à l'épaisseur des parois de Bloch ne peuvent évidemment apparaître que des parois perpendiculaires au plan du film; suivant l'épaisseur ces films sont monodomaines. Si le film est déposé sur un substrat dans un champ magnétique agissant dans le plan du film, il acquiert une anisotropie magnétique *uniaxiale* avec l'axe de facile aimantation dirigé le long du champ. Une telle anisotropie peut être obtenue par recuit des films déjà déposés dans un champ magnétique suffisamment intense ainsi que par dépôt du film par pulvérisation à l'aide d'un faisceau moléculaire dirigé sous un certain angle sur le support.

Considérons de plus près le processus d'aimantation des films. Plaçons un film aimanté à saturation le long de la direction de facile aimantation dans un champ magnétique **H**. Sous l'action de ce champ le vecteur aimantation $\mathbf{J}_m$ du film s'écartera de la direction de facile aimantation d'un certain angle θ (fig. 7.15, *a*). Il en résulte l'apparition dans le film de l'énergie d'anisotropie magnétique U_k. La variation de U_k en fonction de θ est décrite par l'expression

$$U_k = K \sin^2 \theta, \tag{7.30}$$

où K est une constante d'anisotropie.

En plus de l'énergie d'anisotropie (7.30) un film aimanté placé dans un champ magnétique **H** possède encore l'énergie magnétique U_m due à l'interaction entre $\mathbf{J}_m$ et **H**:

$$U_m = -\mu_0 \mathbf{H} \mathbf{J}_m = -\mu_0 H_f J_m \cos \theta - \mu_0 H_d J_m \sin \theta, \tag{7.31}$$

où H_f et H_d sont les composantes du champ extérieur le long de directions de facile et de difficile aimantation respectivement. L'énergie totale du film a pour expression

$$U = U_k + U_m = K \sin^2 \theta - \mu_0 H_f J_m \cos \theta - \mu_0 H_d J_m \sin \theta. \quad (7.32)$$

Supposons qu'un film aimanté suivant la direction de facile

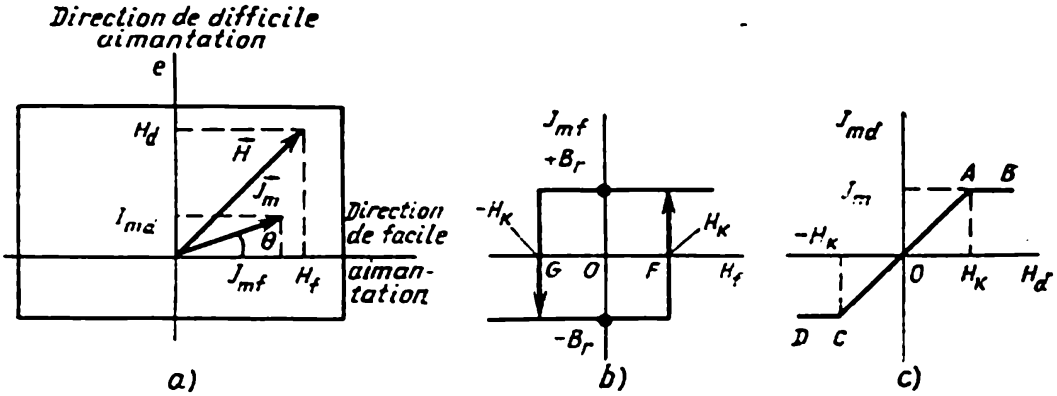

Fig. 7.15. Explication de l'aimantation des films minces.

aimantation ($\theta = 0$) est soumis à un champ extérieur H_f dirigé contre J_m. Alors la variation de l'énergie du film en fonction de l'angle θ est de la forme

$$U = K \sin^2 \theta + \mu_0 H J_m \cos \theta. \quad (7.33)$$

Pour un champ extérieur relativement faible H, cette variation est représentée sur la figure 7.16, *a*. On voit que pour $\theta = 0$ la fonction $U(\theta)$ passe par un minimum (bien que non absolu) de sorte que c'est la conservation de l'orientation initiale du vecteur aimantation J_m (opposée au champ H) qui est énergétiquement avantageuse. On peut s'assurer de l'existence d'un tel minimum en dérivant (7.33) par rapport à θ:

$$\frac{dU}{d\theta} = 2K \sin\theta \cos\theta - \mu_0 H J_m \sin\theta = \sin\theta\,(2K\cos\theta - \mu_0 H J_m). \quad (7.34)$$

Il résulte de (7.34) que pour $\theta = 0$ la dérivée $dU/d\theta = 0$, ce qui constitue, comme on le sait, la condition d'extrémum (dans notre cas de minimum) de la fonction $U(\theta)$.

Dans un champ extérieur suffisamment fort ce sera pour $\theta = 0$ non pas un minimum mais un maximum de la fonction $U(\theta)$ (fig. 7.16, *b*). Le passage du minimum au maximum se produit pour un champ H_k pour lequel la dérivée seconde de la fonction $U(\theta)$ devient nulle pour $\theta = 0$:

$$\left.\frac{d^2U}{d\theta}\right|_{\theta=0} = 2K - \mu_0 H_k J_m = 0. \quad (7.35)$$

Ainsi, dès que H dépasse la valeur $H_k = 2K/(\mu_0 J_m)$, le minimum d'énergie pour $\theta = 0$ disparaît et $\mathbf{J}_m$ s'oriente dans la direction de $\theta = \pi$.

La figure 7.15, *b* montre le cycle d'hystérésis d'un tel film. Pour $H_f > H_k$ (à droite du point F) l'aimantation $\mathbf{J}_m$ est dirigée suivant la direction de facile aimantation (sur la figure 7.15, *b* vers le haut), en se confondant avec la direction de $\mathbf{H}_f$. Lorsque le champ

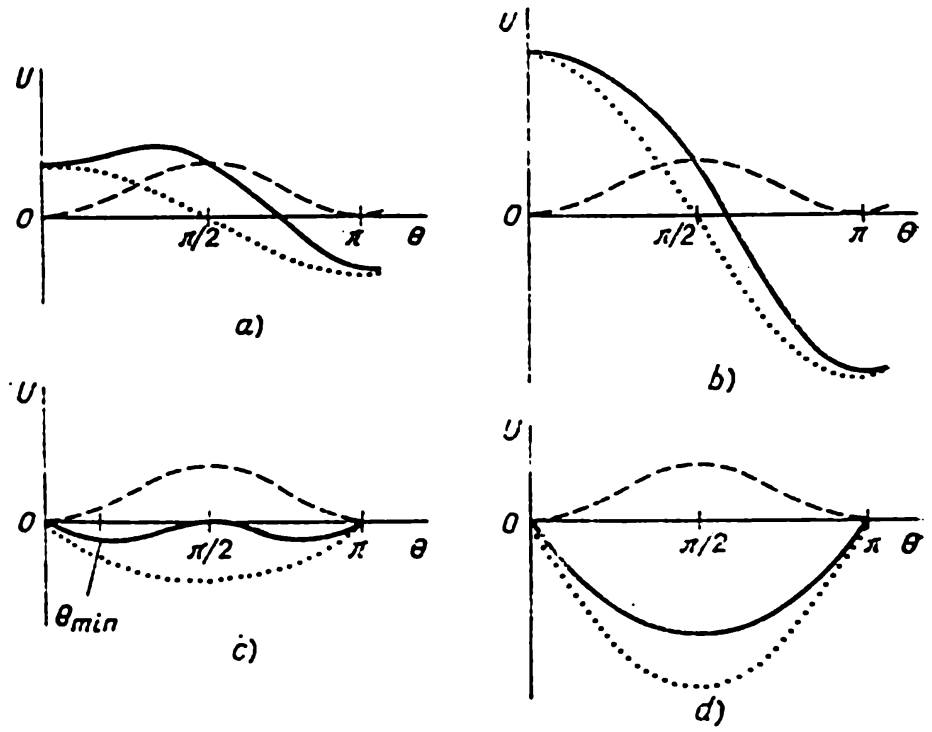

Fig. 7.16. Variation de l'énergie d'un film magnétique mince en fonction de l'orientation du vecteur aimantation $\mathbf{J}_m$ par rapport au champ $\mathbf{H}$:

a — le champ est dirigé suivant la direction de facile aimantation et est inférieur à H_k. La courbe en traits interrompus correspond au terme $K^2 \sin^2\theta$ dans l'expression (7.33); la courbe en pointillé correspond au terme $\mu_0 H J_m \cos\theta$; la courbe en traits pleins correspond à l'énergie totale du film; *b* — idem mais pour $H > H_k$; *c* — le champ $\mathbf{H}$ est perpendiculaire à la direction de facile aimantation et inférieur à H_k. La courbe en traits interrompus correspond au terme $K^2 \sin^2\theta$ dans l'expression (7.36); la courbe en pointillé correspond au terme $\mu_0 H J_m \sin\theta$; *d* — idem mais pour $H > H_k$.

H_f diminue jusqu'à zéro et le champ de sens contraire augmente l'aimantation $\mathbf{J}_m$ reste inchangée jusqu'au point G. En passant par le point G, l'aimantation $\mathbf{J}_m$ change de signe par saut (dirigée vers le bas sur la figure 7.15, *b*) qui reste ensuite inchangé quelque grandes que soient les valeurs de $|H_f| > H_k$.

Une situation analogue se présente aussi dans le cycle inverse: jusqu'au point F y compris le vecteur $\mathbf{J}_m$ est dirigé vers la gauche (sur la figure 7.15, *b* vers le bas) et au point F il bascule par saut vers la droite (sur la figure 7.15, *b* vers le haut). Il en résulte un cycle d'hystérésis rectangulaire à deux états stables: $+B_r$ et $-B_r$.

Maintenant appliquons au film, dont l'aimantation $\mathbf{J}_m$ est dirigée suivant la direction de facile aimantation, un champ extérieur $\mathbf{H}$ dirigé suivant la direction de difficile aimantation. La relation (7.32) prend alors la forme suivante:

$$U = K \sin^2 \theta - \mu_0 H J_m \sin \theta. \tag{7.36}$$

Cette fonction est représentée par la figure 7.16, *c*. On voit que maintenant l'énergie du film est minimale pour $\theta = \theta_{min} > 0$. Pour déterminer θ_{min}, dérivons (7.36) par rapport à θ et annulons la dérivée:

$$\left.\frac{dU}{d\theta}\right|_{\theta=\theta_{min}} = 2K \sin \theta_{min} \cos \theta_{min} - \mu_0 H J_m \cos \theta_m =$$

$$= \cos \theta_{min} (2K \sin \theta_{min} - \mu_0 H J_m) = 0. \tag{7.37}$$

Pour $H < H_k = 2K/(\mu_0 J_m)$ la nullité de $dU/d\theta$ est assurée pour

$$\sin \theta_{min} = \mu_0 H J_m/(2K). \tag{7.38}$$

Pour $H = H_k$ on a $\sin \theta_{min} = 1$ et la position du minimum sera observée pour $\theta_{min} = \pi/2$. Elle sera conservée aussi pour $H > H_k$.

L'aimantation du film suivant la direction de difficile aimantation a pour valeur

$$J_{md} = J_m \sin \theta_{min} = J_m^2 \mu_0 H/(2K). \tag{7.39}$$

Elle croît proportionnellement à H (fig. 7.15, *c*) jusqu'au point A correspondant à $H = H_k$. Pour $H = H_k$ l'aimantation $\mathbf{J}_{md}$ atteint sa valeur maximale

$$J_{md} \mid_{max} = J_m^2 \mu_0 H_k/(2K) = J_m \tag{7.40}$$

et reste inchangée pour $H > H_k$ (partie horizontale AB de la courbe de la figure 7.15, *c*).

Une situation analogue se présente aussi lorsqu'on applique un champ $\mathbf{H}$ de sens contraire (partie OCD de la courbe de la figure 7.15, *c*).

Les processus de renversement de l'aimantation des films minces s'effectuent, comme il a été décrit plus haut, par une application rapide du champ magnétique H au film. On observe alors une rotation *cohérente* simultanée des spins et du vecteur aimantation. Le renversement de l'aimantation peut se faire très vite — pendant un temps de l'ordre de quelques nanosecondes. Si un film aimanté est placé dans un champ magnétique progressivement croissant, dirigé contre le vecteur aimantation du film, il se forme dans le film, pour une certaine intensité du champ, des germes d'aimantation inverse qui croissent à mesure que H augmente, ce qui conduit finalement au renversement de l'aimantation de tout le film. C'est un processus relativement lent qui exige pour sa réalisation quelques microsecondes.

L'intérêt que présentent les films magnétiques réside dans la possibilité de réaliser à leur aide des mémoires pour les calculateurs électroniques qui présentent certains avantages sur les mémoires à ferrites. La figure 7.17 montre un des schémas possibles d'un élément de mémoire à films magnétiques. Cet élément comporte un film *2* de permalloy ou de ferrite déposé par pulvérisation sur un substrat et trois barres métalliques déposées elles aussi par pulvérisation : une barre de position *4*, une barre des nombres *1* et une barre de lecture *3*. L'élément est réalisé de telle sorte que le champ de la barre des nombres soit parallèle à la direction de difficile aimantation et le champ de la barre de position soit parallèle à la direction de facile aimantation. Pour l'écriture de l'information on fait passer une impulsion de courant dans la barre de position, en aimantant le film le long de la direction de facile aimantation. Suivant le sens de cette impulsion le film reste aimanté, après la cessation de l'impulsion, soit à $+B_r$, ce qui correspond à « 1 », soit à $-B_r$, ce qui correspond à « 0 » (v. fig. 7.15, *b*). Pour la lecture une impulsion de courant est appliquée à la barre des nombres. Le champ magnétique de ce courant fait tourner le vecteur aimantation et diminue le flux magnétique le long de la direction de facile aimantation. La diminution de ce flux induit dans la barre de lecture un signal correspondant.

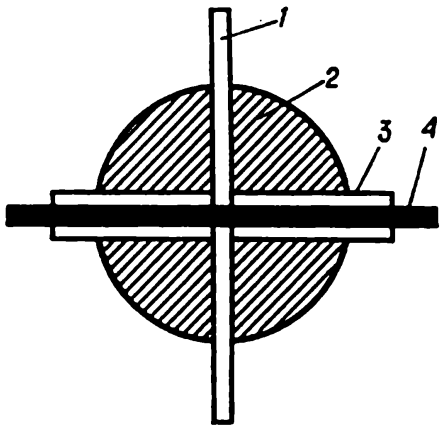

Fig. 7.17. Principe de construction d'une cellule de mémoire à film magnétique plan.

Les mémoires à films magnétiques minces présentent certains avantages sur les mémoires à noyaux en ferrites, ce qui explique les travaux intenses effectués en vue de leur réalisation. Parmi ces avantages on peut citer les suivants : le renversement de l'aimantation des films exige de dépenser une énergie beaucoup plus faible que le renversement de l'aimantation des noyaux ; l'opération pénible de mise en place des noyaux est remplacée par le dépôt par pulvérisation d'un grand nombre de cellules à la fois avec toutes les barres et les connexions nécessaires ; l'utilisation de la rotation cohérente du vecteur aimantation dans les films monodomaines en épaisseur qui permet d'améliorer sensiblement la rapidité de fonctionnement des mémoires.

Domaines magnétiques cylindriques (DMC). Considérons maintenant un film ferromagnétique monocristallin mince (de quelques dixièmes de micromètre à quelques micromètres) comportant un axe

de facile aimantation unique dirigé perpendiculairement au plan du film. Les matériaux utilisés pour la fabrication de tels films sont le plus souvent des grenats uniaxiaux. Les films eux-mêmes sont obtenus par croissance par épitaxie sur des substrats non magnétiques.

A l'état désaimanté, l'aire occupée par les domaines d'une direction de l'aimantation est égale à l'aire des domaines de direction contraire de l'aimantation et les domaines eux-mêmes représentent de minuscules rubans étroits de forme quelconque (fig. 7.18, *a*).

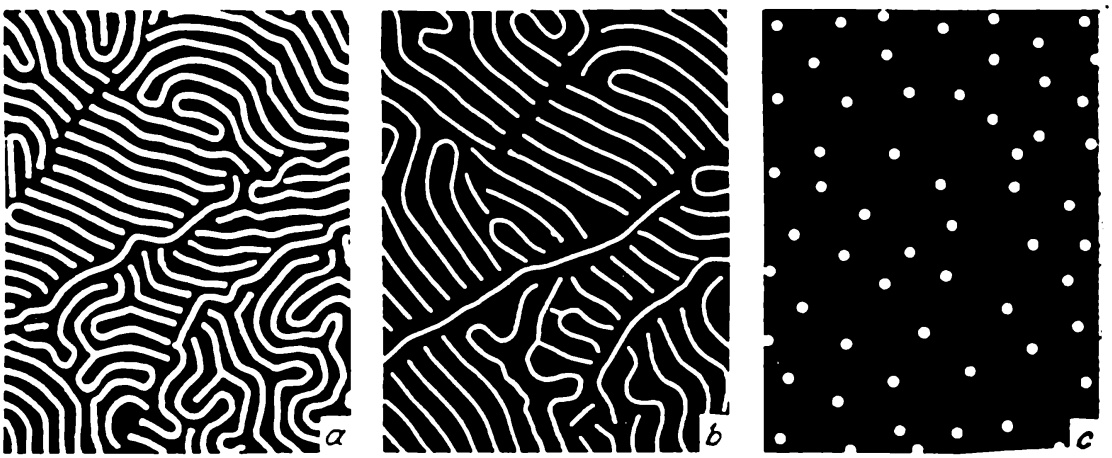

Fig. 7.18. Structure en domaines d'un film à direction de facile aimantation perpendiculaire à la surface:

a — à l'état désaimanté ; *b* — en présence d'un champ de polarisation ; *c* — lorsque le champ de polarisation correspond à l'existence d'un domaine magnétique cylindrique.

Appliquons maintenant à l'échantillon désaimanté un champ magnétique extérieur (champ de polarisation) perpendiculaire au plan du film. Une augmentation progressive de l'intensité du champ provoque, comme toujours lors de l'aimantation d'un ferromagnétique, un déplacement des parois de Bloch et les domaines à aimantation dirigée le long du champ extérieur augmentent au détriment des domaines dont l'aimantation est de sens contraire à celui du champ (fig. 7.18, *b*). Finalement, le champ de polarisation étant suffisamment intense, les domaines à aimantation contraire au champ extérieur disparaissent de sorte que tout le film se trouve aimanté suivant le champ.

Mais il y a une circonstance bien intéressante à signaler : dans une certaine gamme de champs de polarisation inférieurs au champ d'aimantation complète la forme énergétiquement plus avantageuse est une forme cylindrique montrée sur la figure 7.18, *c* et non pas celle des domaines étroits. Aussi, pour une certaine intensité du champ de polarisation, y a-t-il formation de domaines magnétiques

cylindriques (DMC) disposés de façon désordonnée au lieu des domaines en rubans. Ils existent dans le matériau tant que le champ de polarisation n'augmente jusqu'à un *champ* dit *de collapsus* pour lequel les domaines cylindriques devenus très petits, disparaissent totalement.

Si on diminue maintenant le champ de polarisation au-dessous du champ de collapsus, les domaines cylindriques ne réapparaissent pas parce que pour cela la formation de germes est nécessaire, ce qui n'est possible qu'en appliquant au film un champ inverse dont l'intensité est supérieure à l'intensité du champ de collapsus. Si un tel champ est appliqué à tout le film, il s'aimante tout simplement dans le sens inverse. Mais si le champ de polarisation direct est rendu inférieur au champ de collapsus et on applique le champ inverse de façon *localisée* (par exemple, en faisant passer une impulsion de courant par une petite boucle conductrice placée au-dessus du film) un domaine cylindrique prend naissance au-dessous de la boucle et subsiste après la fin de l'impulsion de courant. Ainsi, ce domaine cylindrique est un élément de mémoire dont la présence signifie que la boucle a été parcourue par une impulsion de courant. Dans les films minces faits en matériaux tels que les grenats magnétiquement uniaxiaux, le rayon des domaines cylindriques ne constitue que quelques dixièmes de micromètre de sorte que la densité de stockage de l'information peut être très élevée.

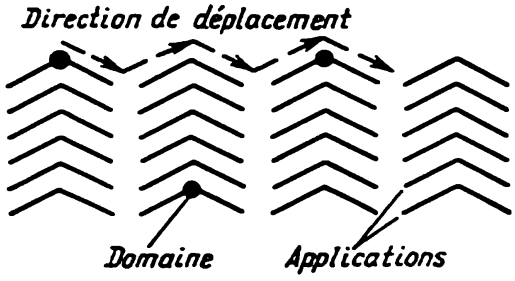

Fig. 7.19. Exemple de structure pour le déplacement du domaine magnétique cylindrique.

Il est remarquable que pour l'écriture de l'information il n'est pas nécessaire de faire défiler le film au-dessous de la boucle d'écriture parce que le domaine cylindrique peut être sorti de dessous de la boucle et déplacé sur le film. On peut le faire par exemple à l'aide d'un champ magnétique tournant extérieur auxiliaire dirigé le long du film si l'on porte sur la surface du film des îlots de forme spéciale (*applications*) en matériau ferromagnétique (généralement permalloy) aimanté le long du film (fig. 7.19).

En laissant ici de côté la mécanique de déplacement des domaines cylindriques au-dessous des applications, indiquons seulement qu'actuellement la dimension d'une cellule de mémoire se détermine par la dimension des éléments d'application et non par celle du domaine cylindrique. En fabriquant des circuits avec une période de près de 3 μm par les méthodes de lithographie électronique, on peut atteindre une densité de stockage de l'information de l'ordre de $\sim 10^7$ bits/cm^2.

La lecture de l'information sur la présence ou l'absence de domaines cylindriques peut se faire en laissant passer à travers le film (qui est d'ordinaire transparent) une lumière polarisée. Le passage de cette lumière à travers le film fait tourner le plan de polarisation dans un sens ou dans l'autre suivant la direction de l'aimantation du film (*effet Faraday*). Une autre méthode d'enregistrement des domaines cylindriques est fondée sur le fait que le passage d'un domaine au-dessous d'une application de permalloy mince fait varier la résistivité électrique de cette application. Pour l'enregistrement des domaines cylindriques on peut aussi utiliser l'effet Hall dans un petit échantillon semiconducteur placé sous le film ainsi que d'autres méthodes.

CHAPITRE 8

PORTEURS DE CHARGES HORS D'ÉQUILIBRE DANS LES SEMICONDUCTEURS

La rapidité de fonctionnement de nombreux appareils à semiconducteurs se détermine par la vitesse de déroulement des processus de formation et de disparition des porteurs de charges libres. Le présent chapitre est consacré à l'étude des principaux facteurs qui contrôlent la cadence de ces processus. On y examine également les phénomènes de diffusion et de dérive des porteurs de charges libres en cas de leur répartition inégale dans le semiconducteur. Ces processus déterminent les paramètres des appareils tels que les transistors, les photodiodes, etc.

§ 8.1. Porteurs de charges en équilibre et hors d'équilibre. Quasi-niveaux de Fermi

Notions sur les porteurs de charges hors d'équilibre. A une température différente de 0 K il se produit dans les semiconducteurs une excitation thermique (*génération*) de porteurs de charges libres. Si ce processus était le seul à se produire, la concentration de porteurs croîtrait indéfiniment avec le temps. Mais il y a un autre processus qui se déroule de pair avec celui de génération, c'est le processus de *recombinaison* : les électrons ayant passé dans la bande de conduction ou aux niveaux accepteurs reviennent dans la bande de valence ou aux niveaux donneurs, ce qui conduit à une diminution de la concentration des porteurs de charges libres. Un *équilibre dynamique* entre ces processus, à toute température, conduit à l'établissement d'une *concentration d'équilibre des porteurs* décrite par les formules (5.13) et (5.16). On dit que ces porteurs sont *en équilibre*.

Les porteurs de charges libres peuvent être produits dans les semiconducteurs non seulement par excitation thermique mais aussi par d'autres procédés : sous l'action de la lumière, sous l'effet des particules ionisantes, par leur injection à travers un contact, etc. L'action de tels agents a pour effet de faire apparaître des porteurs libres additionnels, excédentaires par rapport aux porteurs de charges en équilibre. Désignons par n la concentration totale des électrons et par p la concentration totale des trous. Alors la concentration des porteurs en excès ou *en non-équilibre* sera égale à

$$\Delta n = n - n_0, \quad \Delta p = p - p_0, \tag{8.1}$$

où n_0, p_0 sont les concentrations des porteurs en équilibre.

Si l'intensité de l'agent extérieur reste inchangée, la concentration des porteurs excédentaires croît d'abord bien vite mais ensuite, la vitesse de recombinaison augmentant, sa croissance se ralentit et un état *stationnaire* s'établit finalement pour lequel la vitesse de génération des porteurs est égale à la vitesse de leur recombinaison.

Quasi-niveaux de Fermi. La concentration d'équilibre des électrons et des trous dans un semiconducteur se détermine par le niveau de Fermi d'équilibre et par la température. Elle est décrite par les formules (5.13) et (5.16). Il est essentiel de noter que dans les conditions d'équilibre les électrons et les trous situés dans les bandes et aux niveaux localisés ont un même et seul niveau de Fermi (fig. 8.1, *a*). Les porteurs en équilibre obéissent à la loi d'action de masse (5.36).

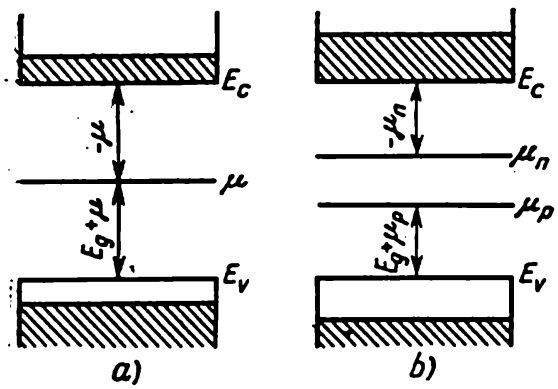

Fig. 8.1. Niveau de Fermi pour des semiconducteurs ne contenant que des porteurs de charges en équilibre (*a*) et quasi-niveaux de Fermi pour des semiconducteurs contenant des électrons et des trous en excès (*b*).

Dans les conditions hors d'équilibre les concentrations des électrons et des trous sont différentes de celles d'équilibre. De (8.1) on tire

$$n = n_0 + \Delta n, \quad p = p_0 + \Delta p. \tag{8.2}$$

Quelque élevés que soient les niveaux dans la bande de conduction auxquels sont excités les électrons sous l'effet de la lumière, des particules ionisantes et d'autres facteurs, ils descendent très vite (pendant $\sim 10^{-11}$ à 10^{-12} s) au niveau inférieur de la bande de conduction et se répartissent suivant les énergies de la même façon que les porteurs en équilibre; quant aux trous hors d'équilibre, ils montent vers le sommet de la bande de valence. Ceci étant, les propriétés des porteurs en excès ne diffèrent pratiquement en rien des propriétés des porteurs en équilibre. Notamment, si l'apparition de porteurs en excès ne trouble pas le caractère non dégénéré du gaz de porteurs libres, pour décrire sa répartition en fonction des énergies on peut utiliser la fonction de répartition d'équilibre (5.12). Il convient seulement de changer la valeur de l'énergie de Fermi μ parce qu'elle détermine le nombre total de porteurs libres dans la bande de conduction et dans la bande de valence qui est devenu maintenant différent. Au lieu de (5.13) et (5.16) il faut écrire

$$n = n_0 + \Delta n = N_c \exp\left[\mu_n/(kT)\right], \tag{8.3}$$

$$p = p_0 + \Delta p = N_v \exp\left(-\frac{E_g+\mu_p}{kT}\right). \tag{8.4}$$

Ici, μ_n et μ_p (que l'on appelle *quasi-niveaux de Fermi*) jouent le rôle des niveaux de Fermi pour les électrons et les trous. Ils ne coïncident ni l'un avec l'autre, ni avec le niveau de Fermi d'équilibre (fig. 8.1, *b*). Les porteurs hors d'équilibre n'obéissent non plus à la loi d'action de masse: $np \neq n_i^2$.

§ 8.2. Durée de vie

Le processus de génération de porteurs est caractérisé par la *vitesse* ou la *cadence de génération* qui est égale au nombre de porteurs (ou de paires de porteurs) excités par seconde dans un volume unité du semiconducteur.

Le processus de recombinaison est caractérisé par la *vitesse de recombinaison* qui s'exprime par le nombre de porteurs (de paires de porteurs) qui se recombinent chaque seconde dans un volume unité du semiconducteur. Chaque porteur de charge libre, qui fait son apparition dans une bande correspondante par suite de la génération, y séjourne, avant de se recombiner, en moyenne un certain temps qu'on appelle *temps* (ou *durée*) *de vie* τ. La vitesse de recombinaison, la concentration des porteurs et leur durée de vie sont liées entre elles par les relations

$$R_n = n/\tau_n, \quad R_p = p/\tau_p, \tag{8.5}$$

où R_n, R_p et τ_n, τ_p sont les vitesses de recombinaison et les durées de vie des électrons et des trous respectivement.

Si les électrons et les trous apparaissent dans le semiconducteur par paires, ils se recombinent également par paires et, à l'état stationnaire, leurs vitesses de recombinaison sont égales l'une à l'autre. Alors

$$n/\tau_n = p/\tau_p. \tag{8.6}$$

Ordinairement $n \neq p$ et donc $\tau_n \neq \tau_p$.

Comme le montre (8.6), la durée de vie des porteurs majoritaires est plus grande que celle des porteurs minoritaires. En outre, les durées de vie elles-mêmes varient en fonction de la concentration des porteurs hors d'équilibre Δn et Δp. Supposons que $\Delta n = \Delta p$, ce qui est souvent le cas. Alors, le rapprochement entre (8.2) et (8.6) donne

$$\frac{n_0+\Delta n}{\tau_n} = \frac{p_0+\Delta n}{\tau_p}. \tag{8.7}$$

Si $n_0 \neq p_0$ cette égalité ne peut être vérifiée pour des valeurs différentes de Δn que dans le cas où τ_n et τ_p varient en fonction de Δn. Pour éviter des complications qui en résultent, on a convenu de ranger les porteurs en équilibre n_0 et p_0 dans un groupe distinct avec les durées de vie τ_n et τ_p qui assurent l'égalité des vitesses de recombinaison d'équilibre

$$R_0 = n_0/\tau_n = p_0/\tau_p, \tag{8.8}$$

et de considérer que les durées de vie τ_n et τ_p, et par conséquent les vitesses de recombinaison des porteurs en équilibre, ne dépendent pas de Δn. Pour ce qui est des porteurs hors d'équilibre, on leur attribue leur durée de vie τ qui détermine leur vitesse de recombinaison

$$R = -\frac{d(\Delta n)}{dt} = +\frac{\Delta n}{\tau}. \tag{8.9}$$

Le signe moins indique qu'au cours de la recombinaison la concentration des porteurs diminue.

Supposons que dans le semiconducteur des porteurs excédentaires $\Delta n_0 = \Delta p_0$ sont excités sous l'action de la lumière. Après l'élimination de la lumière ces porteurs se recombinent de sorte que leur concentration diminue progressivement. En intégrant l'équation (8.9), on obtient

$$\Delta n = \Delta n_0 \exp(-t/\tau). \tag{8.10}$$

L'équation (8.10) montre que $\Delta n = \Delta n_0/e$ pour $t = \tau$.

Ainsi, *la durée de vie moyenne des porteurs de charge excédentaires est égale au temps τ au cours duquel leur concentration devient, par suite de la recombinaison, e fois plus faible.*

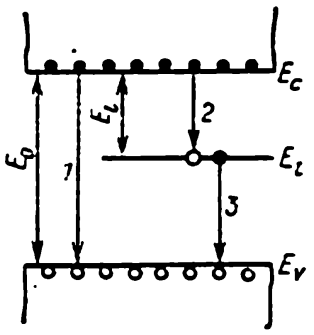

Fig. 8.2. Schéma de la recombinaison bande à bande et de la recombinaison par un niveau localisé dans la bande interdite.

Le processus de transition de l'électron de la bande de conduction dans la bande de valence lors de la recombinaison peut se faire directement à travers toute la bande interdite E_g comme l'indique la flèche *1* de la figure 8.2 ou d'abord à un niveau d'impureté E_t (flèche *2*) et ensuite du niveau d'impureté dans la bande de valence (flèche *3*). Dans le premier cas la recombinaison est dite *recombinaison bande à bande* et dans le second cas *recombinaison par l'intermédiaire du niveau d'impureté*. Examinons-les avec plus de détails.

Recombinaison bande à bande. Dans les conditions d'équilibre thermique, la vitesse de recombinaison bande à bande R_0 est proportionnelle à la concentration des électrons n_0 et à celle des trous p_0 et elle doit être égale à la vitesse de génération g_0 :

$$R_0 = \gamma n_0 p_0 = \gamma n_i^2 = g_0. \tag{8.11}$$

Le coefficient de proportionnalité

$$\gamma = g_0/n_i^2 \tag{8.12}$$

est appelé *coefficient de recombinaison.*

Si le semiconducteur est soumis à l'action d'un facteur extérieur, qui fait augmenter de g la vitesse de génération de porteurs, la vitesse de leur recombinaison augmentera de R. La concentration stationnaire des électrons $n = n_0 + \Delta n$ et des trous $p = p_0 + \Delta p$ s'établira lorsque la vitesse totale de recombinaison $R_0 + R$ deviendra égale à la vitesse totale de génération $g_0 + g$:

$$R_0 + R = \gamma (n_0 + \Delta n)(p_0 + \Delta p) = g_0 + g. \qquad (8.13)$$

En y portant la valeur de γ tirée de (8.10), on obtient

$$g_0 + g = (g_0/n_i^2)(n_0 p_0 + n_0 \Delta p + p_0 \Delta n + \Delta n \Delta p). \qquad (8.14)$$

Pour un niveau bas d'excitation du semiconducteur par un facteur extérieur, lorsque Δn, $\Delta p \ll (n_0 + p_0)$, le dernier terme dans le second membre de (8.14) peut être négligé; en outre, compte tenu que $\Delta n = \Delta p$, $n_0 p_0 = n_i^2$, on trouve

$$R = g = (g_0/n_i^2)(n_0 + p_0)\,\Delta n. \qquad (8.15)$$

Il résulte de (8.15) que dans ces conditions *la vitesse de recombinaison des porteurs de charges excédentaires est proportionnelle à leur concentration*. Une telle recombinaison est dite *linéaire*. La durée de vie des porteurs hors d'équilibre lors de la recombinaison linéaire a pour valeur

$$\tau = \Delta n/R = \Delta n/g = n_i^2/[g_0 (n_0 + p_0)]. \qquad (8.16)$$

Elle ne dépend pas de l'intensité de l'excitation et se détermine par le degré de dopage du semiconducteur (n_0 ou p_0), en décroissant en raison inverse de la concentration de l'impureté de dopage. En effet, pour un semiconducteur de type n contenant une impureté donatrice N_d, on a dans le domaine de déplétion $n_0 = n_{n0} \approx N_d \gg p_0 = p_{n0}$ et $\tau_n = \frac{n_i^2}{g_0} \cdot \frac{1}{n_{n0}} \approx \frac{n_i^2}{g_0} \cdot \frac{1}{N_d}$; pour un semiconducteur de type p $\tau_p = \frac{n_i^2}{g_0} \cdot \frac{1}{p_{p0}} \approx \frac{n_i^2}{g_0} \cdot \frac{1}{N_a}$. La durée de vie τ a sa valeur maximale dans les semiconducteurs intrinsèques pour lesquels $n_0 = p_0 = n_i$. Dans ce cas, (8.16) prend la forme

$$\tau_i = n_i/(2g_0). \qquad (8.17)$$

Pour un niveau d'excitation élevé quand $\Delta n = \Delta p \gg (n_0 + p_0)$, le dernier terme du second membre de (8.14) devient prédominant de sorte que

$$R = g = \frac{g_0}{n_i^2}(\Delta n)^2, \qquad (8.18)$$

c est-à-dire que *la vitesse de recombinaison des porteurs excédentaires devient proportionnelle au carré de leur concentration*. Une telle recombinaison est dite *quadratique*. La durée de vie des porteurs hors

d'équilibre lors de la recombinaison quadratique

$$\tau = \frac{\Delta n}{R} = \frac{\Delta n}{g} = \frac{n_i^2}{g_0} \frac{1}{\Delta n} \tag{8.19}$$

est inversement proportionnelle à leur concentration.

Lors de la recombinaison bande à bande l'énergie peut se dégager soit sous la forme d'un quantum de lumière $h\nu$, soit sous la forme de la chaleur (phonons). Dans le premier cas elle est dite *radiative* et dans le second cas *non radiative*. Comme le montrent les calculs et les expériences, la recombinaison bande à bande radiative peut avoir une importance substantielle pour les semiconducteurs à bande interdite étroite aux températures relativement hautes (ordinaire et plus élévée). Quant aux semiconducteurs à bande interdite large, le mécanisme principal, responsable de la recombinaison, y est la recombinaison non radiative par les niveaux d'impureté. Toutefois, dans certaines conditions on peut atteindre dans ces semiconducteurs eux aussi un niveau relativement haut de recombinaison radiative. Comme il résulte de (8.16) et (8.19), cela est favorisé notamment par l'augmentation de la concentration des porteurs excédentaires dans le semiconducteur et l'augmentation du degré de son dopage.

Recombinaison par les niveaux localisés. Le mécanisme de la recombinaison à travers les niveaux localisés consiste en ce que l'électron passe de la bande de conduction d'abord à un niveau localisé dans la bande interdite et ensuite à un niveau libre dans la bande de valence (fig. 8.2). Une telle recombinaison en deux étapes est souvent beaucoup plus probable que la recombinaison directe d'un électron et d'un trou examinée plus haut.

Les niveaux localisés dans la bande interdite peuvent constituer des centres efficaces de recombinaison s'ils sont situés loin du niveau inférieur de la bande de conduction et du sommet de la bande de valence; dans le cas contraire ils jouent le rôle des niveaux de collage parce que le porteur capté par eux est réinjecté, au bout d'un certain temps, dans sa bande. Les centres de recombinaison sont souvent appelés *pièges* et le processus de passage de l'électron de la bande de conduction à un niveau libre de recombinaison est appelé *capture d'un électron* par un piège. D'une manière analogue, le passage d'un électron d'un piège à un niveau libre dans la bande de valence avec libération du niveau de piège s'appelle *capture d'un trou* par un piège.

Examinons ces processus de façon un peu plus détaillée. Séparons dans le semiconducteur une couche à faces parallèles d'épaisseur dx, perpendiculaire à la direction de mouvement des électrons (fig. 8.3, *a*). Désignons par N_p la concentration des pièges. En considérant la capture d'un électron, on peut assimiler chaque piège à une sphère de rayon r. En pénétrant dans cette sphère, l'électron subit une collision et est capturé par le piège. Si l'on effectue, par la pensée,

la radioscopie de la couche séparée, on verra que les pièges constituent une sorte de disques de surface $A_n = \pi r^2$ (fig. 8.3, *b*). Si un électron passant par la couche tombe dans un de tels disques il est capturé par le piège. C'est pour cette raison que A_n est appelée *section efficace de capture d'un électron.*

Désignons par S la surface de la couche séparée. Le volume de cette couche est égal à $S\,dx$ et le nombre de pièges qu'il contient est

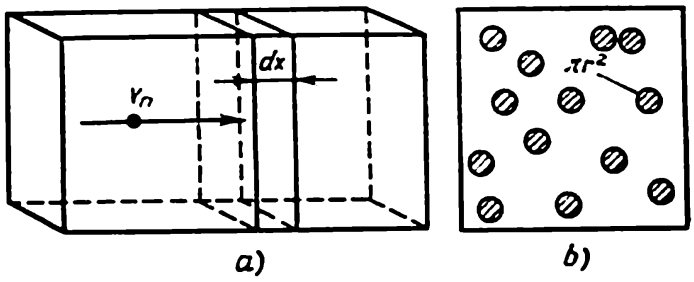

Fig. 8.3. Etablissement de la relation pour la probabilité de capture de porteurs par des pièges.

$N_{p.lib}S\mathrm{d}x$, où $N_{p.lib}$ est la concentration des pièges non occupés par les électrons (seuls ces pièges sont capables de capturer les électrons). La surface totale de capture de ces pièges a pour valeur

$$S_{capt} = A_n N_{p.lib} S\mathrm{d}x.$$

La probabilité pour qu'un électron traversant la couche subisse une collision avec un piège est évidemment égale au rapport de S_{capt} à S :

$$\mathrm{d}W_n = S_{capt}/S = A_n N_{p.lib}\mathrm{d}x.$$

Comme $\mathrm{d}x = v_n \mathrm{d}t$, où v_n est la vitesse de l'électron, on a

$$\mathrm{d}W_n = A_n N_{p.lib}\, v_n \mathrm{d}t.$$

En divisant par $\mathrm{d}t$ le premier et le second membre de cette relation, on obtient

$$\mathcal{P}_n = \frac{\mathrm{d}W_n}{\mathrm{d}t} = A_n N_{p.lib}\, v_n, \tag{8.20}$$

où $\mathcal{P}_n$ est la probabilité de capture d'un électron par un piège rapportée à l'unité de temps.

Une expression analogue peut être obtenue pour la probabilité de capture d'un trou :

$$\mathcal{P}_p = A_p N_{p.occ} v_p, \tag{8.21}$$

où A_p est la section efficace de capture d'un trou par un piège ; $N_{p.\,occ}$ la concentration des pièges occupés par les électrons (libres pour la capture de trous) ; v_p la vitesse moyenne des trous.

Il convient de noter que dans le cas général $A_n \neq A_p$; les valeurs de ces sections efficaces sont généralement comprises dans les limites de 10^{-17} à 10^{-25} m².

En plus des processus de capture des électrons et des trous par des pièges, il convient de tenir compte aussi des processus inverses, à savoir de l'excitation thermique des électrons situés aux niveaux des pièges dans la bande de conduction et de la bande de valence aux niveaux libres des pièges. Ce dernier processus peut être considéré comme une expulsion thermique des trous des niveaux des pièges dans la bande de valence. Dans le diagramme énergétique (v. fig. 8.2) ces processus pourraient être représentés par des flèches contraires aux flèches *2* et *3*.

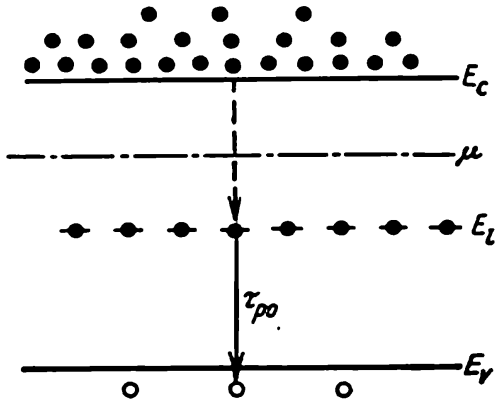

Fig. 8.4. Schéma de la recombinaison par un niveau des pièges dans un semiconducteur de type *n* manifeste.

Dans certains cas particuliers de la recombinaison par les pièges, il est facile d'obtenir des expressions pour la durée de vie des porteurs de courant. Considérons un semiconducteur qui ne comporte qu'un seul niveau des pièges situé près du milieu de la bande interdite. Dans ce cas, dans un semiconducteur de type *n* assez fortement dopé, le niveau de Fermi μ se situe, dans les conditions d'équilibre thermique, de plusieurs kT plus haut que le niveau des pièges E_l et tous les pièges sont occupés par les électrons (fig. 8.4). De tels pièges ne peuvent pas capturer des électrons de la bande de conduction mais ils sont capables de relâcher leurs électrons qui passent dans la bande de valence, c'est-à-dire de capturer un trou (flèche en traits pleins de la figure 8.4). Le temps moyen nécessaire pour la capture d'un trou par n'importe lequel des pièges est égal à

$$\tau_{p0} = 1/\mathcal{P}_{p0} = 1/(A_p N_p v_p), \tag{8.22}$$

où $\mathcal{P}_{p0}$ désigne la probabilité de capture d'un trou par des pièges rapportée à l'unité de temps (8.21) dans les conditions lorsque tous les pièges sont occupés par les électrons, c'est-à-dire lorsque $N_{p.occ} = N_p$.

Dès que le piège devient libre, il sera occupé immédiatement par un des électrons de la bande de conduction qui sont très nombreux dans un semiconducteur de type *n* fortement dopé (flèche en traits interrompus de la figure 8.4) et l'acte de recombinaison sera achevé.

Ainsi, l'expression (8.22) détermine la durée de vie des porteurs minoritaires — des trous — dans les conditions d'équilibre thermique. Pour un faible niveau d'excitation du semiconducteur lors-

qu'aux électrons et aux trous en équilibre s'ajoutent des porteurs de charges hors d'équilibre, les pièges sont comme précédemment occupés par les électrons et les conditions de capture des trous par ces pièges restent inchangées. Cela signifie que la formule (8.22) détermine aussi la durée de vie des trous hors d'équilibre, c'est-à-dire tout simplement la durée de vie des porteurs de charges hors d'équilibre.

En opérant de façon identique, on peut obtenir une expression pour la durée de vie des porteurs de charges minoritaires et hors d'équilibre dans un semiconducteur de type p fortement dopé:

$$\tau_{n0} = 1/\mathcal{P}_{n0} = 1/(A_n N_p v_n). \tag{8.23}$$

Dans le cas général, lorsque le degré de dopage du semiconducteur n'est pas trop élevé, l'étude des processus de génération-recombinaison peut s'avérer assez fastidieuse. Hall, Schockley et Read ont effectué une analyse correspondante pour un seul niveau des pièges et ont obtenu pour la vitesse de recombinaison des porteurs hors d'équilibre, dans les conditions stationnaires, l'expression suivante:

$$R = \frac{(n_0+\Delta n)(p_0+\Delta p) - n_0 p_0}{(n_0+\Delta n+n_p)\tau_{p0} + (p_0+\Delta p+p_p)\tau_{n0}}, \tag{8.24}$$

où n_0, p_0 sont les concentrations d'équilibre des électrons dans la bande de conduction et des trous dans la bande de valence; n_p, p_p les concentrations d'équilibre des électrons dans la bande de conduction et des trous dans la bande de valence lorsque le niveau de Fermi coïncide avec le niveau des pièges E_p, et τ_{p0} et τ_{n0} sont déterminés par les expressions (8.22) et (8.23) respectivement.

Dans le cas de la recombinaison à travers les pièges Δn peut ne pas coïncider avec Δp parce qu'une partie d'électrons ou de trous hors d'équilibre peut se trouver dans des pièges en modifiant le degré de leur occupation. Toutefois, si les pièges ne sont pas trop nombreux et le degré d'excitation du semiconducteur n'est pas trop petit de sorte qu'est satisfaite la condition $N_p \ll \Delta n$, Δp, on a $\Delta n \approx \approx \Delta p$ et la formule (8.24) prend alors la forme

$$R = \frac{(n_0+p_0)\Delta n + (\Delta n)^2}{(n_0+n_p+\Delta n)\tau_{p0} + (p_0+p_p+\Delta n)\tau_{n0}}. \tag{8.25}$$

Enfin, si les conditions $\Delta n \ll (n_0 + n_p)$, $(p_0 + p_p)$, $(n_0 + p_0)$ sont satisfaites toutes à la fois, l'expression (8.25) devient

$$R = \frac{(n_0+p_0)\Delta n}{(n_0+n_p)\tau_{p0} + (p_0+p_p)\tau_{n0}}. \tag{8.26}$$

Alors, en faisant usage de la relation (8.9), on peut déterminer la durée de vie des porteurs hors d'équilibre

$$\tau = \frac{\Delta n}{R} = \frac{n_0+n_p}{n_0+p_0}\tau_{p0} + \frac{p_0+p_p}{n_0+p_0}\tau_{n0}. \tag{8.27}$$

En cas d'un fort dopage par exemple pour un semiconducteur de type n, lorsque $n_0 \gg n_p$, p_p, p_0, la relation (8.27) donne $\tau = \tau_{p0}$. De façon analogue, pour un semiconducteur de type p, $\tau = \tau_{n0}$. Cette conclusion a déjà été faite plus haut [v. (8.22) et (8.23)]. Pour de tels semiconducteurs la durée de vie ne dépend pas du degré de dopage par des donneurs ou par des accepteurs de faible énergie, elle se détermine uniquement par la concentration des pièges N_p et par leur type (par les grandeurs A_n et A_p). La figure 8.5 montre les courbes de variation de la durée de vie en fonction du degré de dopage

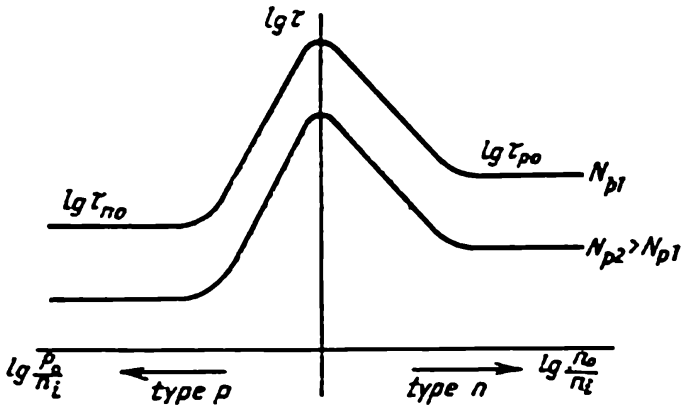

Fig. 8.5. Variation de la durée de vie des porteurs hors d'équilibre en fonction de la concentration des porteurs de courant majoritaires.

d'un semiconducteur pour deux concentrations des pièges (pour fixer les idées on a pris $\tau_{n0} < \tau_{p0}$). C'est dans un semiconducteur intrinsèque que la durée de vie atteint sa valeur maximale.

Dans le cas d'un très haut niveau d'excitation les expressions (8.25) et (8.9) donnent

$$\tau = \tau_{p0} + \tau_{n0}. \tag{8.28}$$

Si plusieurs canaux de recombinaison, indépendants l'un de l'autre, sont possibles, la probabilité de recombinaison doit être égale à la somme des probabilités de recombinaison par des canaux distincts et la durée de vie effective τ doit être déterminée par la relation suivante:

$$\frac{1}{\tau} = \sum_k \frac{1}{\tau_k}, \tag{8.29}$$

où τ_k est la durée de vie déterminée par le k-ième canal de recombinaison. La durée de vie effective se détermine ainsi par la plus petite de ces durées de vie « particulières ».

§ 8.3. Equation de continuité

La concentration des porteurs de charges libres peut varier dans une région ou dans une autre non seulement par suite de leur génération ou leur recombinaison mais également par suite de leur déplacement dans le cristal, c'est-à-dire par suite du passage des courants. Considérons par exemple comment varie le nombre de trous dans un certain élément de volume du cristal $dV = dx dy dz$ (fig. 8.6) par suite du passage des courants de trous de densité $\mathbf{i}_p = \mathbf{i}_p(x, y, z)$. A travers la face *1* de surface $dy\,dz$, pendant un petit temps dt, entrera dans l'élément de volume considéré une charge $i_{px}(x, y, z)\,dy\,dz\,dt$,

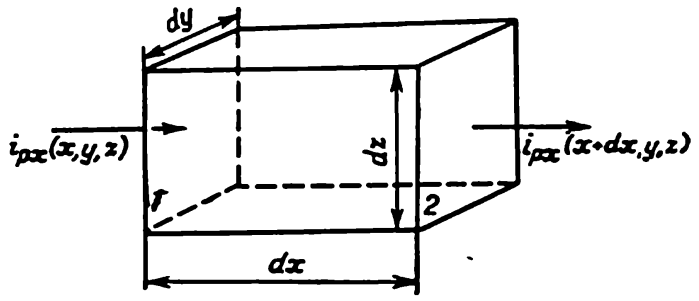

Fig. 8.6. Etablissement de l'équation de continuité.

c'est-à-dire un nombre dP_1 de trous égal à

$$dP_1 = \frac{1}{e}\, i_{px}(x, y, z)\, dy\, dz\, dt.$$

Pendant le même temps, dP_2 trous quitteront le volume considéré par la face *2*:

$$dP_2 = \frac{1}{e}\, i_{px}(x + dx, y, z)\, dy\, dz\, dt.$$

La variation de la concentration des trous au voisinage du point de coordonnées x, y, z sera obtenue si on retranche dP_2 de dP_1 et divise la différence par le volume dV. Ainsi il vient

$$dp = \frac{1}{e}\,[i_{px}(x, y, z) - i_{px}(x + dx, y, z)]\, dt/dx = -\frac{1}{e}\frac{\partial i_{px}}{\partial x}\, dt.$$

Des contributions analogues seront apportées à la variation de la concentration des trous par des composantes du courant de trous suivant les axes des y et des z. Finalement on obtient

$$\left(\frac{\partial p}{\partial t}\right)_{i_p} = -\frac{1}{e}\left(\frac{\partial i_{px}}{\partial x} + \frac{\partial i_{py}}{\partial y} + \frac{\partial i_{pz}}{\partial z}\right) = -\frac{1}{e}\,\mathrm{div}\,\mathbf{i}_p, \qquad (8.30)$$

où $(\partial p/\partial t)_{i_p}$ désigne la variation de la concentration des trous par suite du passage des courants de trous.

La variation de la concentration des électrons a pour expression

$$\left(\frac{\partial n}{\partial t}\right)_{i_n} = \frac{1}{e} \operatorname{div} \mathbf{i}_n. \tag{8.31}$$

Maintenant on peut écrire pour la variation des concentrations une expression tenant compte de tous les facteurs :

$$\begin{cases} \dfrac{\partial p}{\partial t} = g_p - R_p - \dfrac{1}{e} \operatorname{div} \mathbf{i}_p ; \\ \dfrac{\partial n}{\partial t} = g_n - R_n + \dfrac{1}{e} \operatorname{div} \mathbf{i}_n. \end{cases} \tag{8.32}$$

Ces relations portent le nom d'*équations de continuité*. Comme les valeurs d'équilibre des vitesses de génération et de recombinaison (g_0 et R_0) sont égales, on peut dans (8.32) tenir compte seulement des vitesses de génération des porteurs de charge hors d'équilibre. De plus, on peut profiter des (8.2) et (8.9) pour mettre l'équation de continuité sous une forme plus commode :

$$\begin{cases} \dfrac{\partial (\Delta p)}{\partial t} = g - \dfrac{\Delta p}{\tau_p} - \dfrac{1}{e} \operatorname{div} \mathbf{i}_p ; \\ \dfrac{\partial (\Delta n)}{\partial t} = g - \dfrac{\Delta n}{\tau_n} + \dfrac{1}{e} \operatorname{div} \mathbf{i}_n. \end{cases} \tag{8.33}$$

Ici, g sont les vitesses de la génération additionnelle des électrons et des trous (on suppose que la génération est de bande à bande) et τ_n, τ_p les durées de vie des électrons et des trous hors d'équilibre respectivement.

§ 8.4. Diffusion conjointe des porteurs hors d'équilibre

Les courants d'électrons et de trous dans les semiconducteurs prennent naissance par suite de la dérive des porteurs de charges libres dans un champ électrique et de leur diffusion. Les densités des courants de dérive se déterminent par les expressions

$$\mathbf{i}_{n\,der} = enu_n\mathscr{E}, \quad \mathbf{i}_{p\,der} = epu_p\mathscr{E}, \tag{8.34}$$

où u_n et u_p sont les mobilités des électrons et des trous respectivement. Quant aux densités des courants de diffusion, elles ont pour expression

$$\mathbf{i}_{n\,dif} = eD_n \operatorname{grad} n ; \quad \mathbf{i}_{p\,dif} = -eD_p \operatorname{grad} p, \tag{8.35}$$

où D_n et D_p sont les coefficients de diffusion des électrons et des trous respectivement. Les densités totales des courants d'électrons et de trous sont égales à la somme de leurs composantes de dérive et de diffusion :

$$\mathbf{i}_n = \mathbf{i}_{n\,der} + \mathbf{i}_{n\,dif} ; \quad \mathbf{i}_p = \mathbf{i}_{p\,der} + \mathbf{i}_{p\,dif}. \tag{8.36}$$

Les coefficients de diffusion des porteurs de charges sont liés à leurs mobilités. Pour des gaz non dégénérés de porteurs cette liaison est établie par la *relation d'Einstein*:

$$D = \frac{kT}{e} u. \tag{8.37}$$

Dans le cas général, les mobilités et donc les coefficients de diffusion des électrons et des trous sont différents. C'est ainsi que dans le germanium et le silicium la mobilité des électrons est à peu près deux fois celle des trous.

Le champ électrique $\mathscr{E}$ figurant dans (8.34) n'est pas obligatoirement produit par des charges extérieures, il peut prendre naissance aussi dans le cas où l'électroneutralité du semiconducteur lui-même est troublée. Notamment, si le semiconducteur n'est soumis à aucune différence de potentiel extérieure, mais dans une quelconque de ses régions sont générés (inégalement suivant l'échantillon) des porteurs de charges hors d'équilibre, un champ électrique « incorporé » naît automatiquement qui fait équilibre aux courants d'électrons et de trous.

Examinons de plus près ce processus. Supposons qu'une lumière d'énergie des quanta $h\nu > E_g$ tombe sur la surface d'un semiconducteur. Une telle lumière est fortement absorbée dans le semiconducteur en y faisant naître des paires électron-trou. Comme les porteurs hors d'équilibre ne sont pas immobiles mais participent, de même que les porteurs en équilibre, à l'agitation thermique désordonnée, ils diffusent vers les régions profondes du semiconducteur et, chemin faisant, se recombinent. Le coefficient de diffusion des électrons D_n et le coefficient de diffusion des trous D_p sont différents (d'ordinaire $D_n > D_p$) et il semblerait que la concentration des électrons Δn et celle des trous Δp devraient diminuer à mesure qu'on s'éloigne de la surface, comme le montre la figure 8.7, *a*. Toutefois, comme il est visible sur cette figure, dans ce cas il y a apparition de charges d'espace et donc d'un champ électrique. Mais si un champ électrique naît dans un semiconducteur, il se produit une dérive de porteurs dans ce champ qui tend à annuler les charges d'espace, à égaliser les concentrations des porteurs excédentaires dans chaque point du semiconducteur. Il en résulte que les porteurs hors d'équilibre se répartissent comme l'indique la figure 8.7, *b* comme s'ils diffusaient avec un certain coefficient effectif de diffusion D_{eff}, le même pour les électrons et pour les trous. Proposons-nous de déterminer ce coefficient de diffusion effectif. Dans les conditions stationnaires, la densité totale des courants d'électrons et de trous dans chaque section du semiconducteur doit être égale à zéro:

$$i_n + i_p = 0. \tag{8.38}$$

En combinant (8.34) et (8.35) et en tenant compte du fait que la diffusion et la dérive s'effectuent dans le cas considéré suivant une même direction — le long de l'axe des x — récrivons (8.38) sous la forme

$$enu_n \mathscr{E} + eD_n \frac{dn}{dx} + epu_p \mathscr{E} - eD_p \frac{dp}{dx} = 0. \qquad (8.39)$$

Ici $\mathscr{E}$ est l'intensité du champ « incorporé » qui apparaît automatiquement par suite d'une différence entre les coefficients de diffusion des électrons et des trous. Dans (8.39) les dérivées de n et p peuvent être remplacées par les dérivées de Δn et Δp. En outre nous

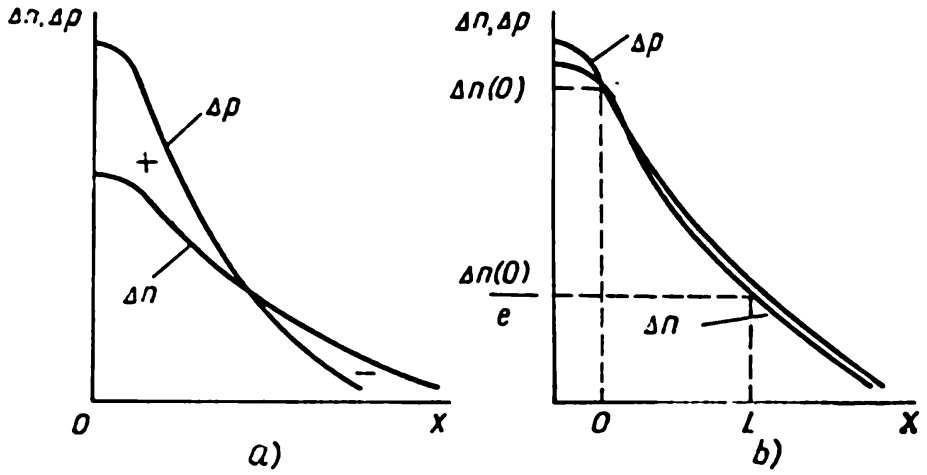

Fig. 8.7. Explication de la cause d'une diffusion conjointe des porteurs hors d'équilibre.

admettrons que pratiquement $\Delta n = \Delta p$. Alors, on tire de (8.39)

$$\mathscr{E} = (D_p - D_n) \frac{d(\Delta n)}{dx} \frac{1}{(nu_n + pu_p)}. \qquad (8.40)$$

Ceci étant, la composante électronique i_n de la densité de courant peut s'écrire sous la forme

$$i_n = enu_n (D_p - D_n) \frac{d(\Delta n)}{dx} \frac{1}{(nu_n + pu_p)} + eD_n \frac{d(\Delta n)}{dx} =$$
$$= e \frac{(nu_n D_p + pu_p D_n)}{(nu_n + pu_p)} \frac{d(\Delta n)}{dx}. \qquad (8.41)$$

L'expression obtenue peut être interprétée de façon formelle comme donnant le courant dû uniquement à la diffusion [comparer à (8.35)], si le facteur devant d (Δn) dx est pris pour coefficient de diffusion effectif des électrons hors d'équilibre. Une expression absolument identique est obtenue pour le coefficient de diffusion effec-

tif des trous hors d'équilibre. En remplaçant dans (8.41) les mobilités par les coefficients de diffusion conformément à la relation d'Einstein (8.37), on obtient

$$D_{eff} = (n+p) \Big/ \left(\frac{n}{D_p} + \frac{p}{D_n} \right). \tag{8.42}$$

Pour un semiconducteur intrinsèque $D_{eff} = 2D_n D_p/(D_n + D_p)$, c'est-à-dire que son coefficient de diffusion effectif a une valeur intermédiaire entre celles des coefficients de diffusion des électrons er des trous. Quant aux semiconducteurs extrinsèques manifestes, leut coefficient de diffusion effectif est égal au coefficient de diffusion des porteurs *minoritaires*. C'est ainsi que $D_{eff} = D_p$ pour $n \gg p$ et $D_{eff} = D_n$ pour $p \gg n$. Dans ces cas, la charge d'espace des porteurs minoritaires est facilement compensée par de nombreux porteurs majoritaires qui « poursuivent » constamment les porteurs minoritaires en diffusion.

Déterminons maintenant la répartition stationnaire ($\partial(\Delta n)/\partial t = 0$) des porteurs hors d'équilibre derrière la couche dans laquelle la lumière tombante provoque une génération appréciable de paires électron-trou. Il est commode de prendre la coordonnée de l'extrémité de cette couche pour zéro (origine). L'équation de continuité pour $x > 0$, où $g = 0$, prend alors la forme suivante :

$$-\frac{\Delta n}{\tau} + \frac{1}{e}\frac{di_n}{dx} = -\frac{\Delta n}{\tau} + D_{eff}\frac{d^2(\Delta n)}{dx^2} = 0. \tag{8.43}$$

Ici τ est la durée de vie des porteurs hors d'équilibre. La solution de cette équation est de la forme

$$\Delta n(x) = \Delta n(0)\, e^{-x/\sqrt{D_{eff}\tau}} = \Delta n(0)\, e^{-x/L}, \tag{8.44}$$

où

$$L = \sqrt{D_{eff}\tau} \tag{8.45}$$

est une *longueur* dite *de diffusion*. Elle détermine la valeur moyenne de la profondeur de pénétration des porteurs hors d'équilibre depuis des régions où s'effectue la génération de paires électron-trou.

Notons que si les porteurs injectés dans le semiconducteur à travers le plan $x = 0$ étaient uniquement des porteurs minoritaires, une charge d'espace y prendrait naissance. Pour compenser cette charge, des porteurs majoritaires viendraient de la profondeur du semiconducteur de sorte que la répartition des électrons et des trous excédentaires pour $x > 0$ serait la même que celle obtenue plus haut.

Faisons maintenant quelques évaluations numériques pour justifier les arguments avancés. Considérons par exemple le silicium de type n avec une concentration des électrons $n_0 = 10^{16}$ cm^{-3}. La

concentration des porteurs dans un semiconducteur intrinsèque à la température ordinaire $n_i \approx 10^{10}$ cm^{-3}. Alors, suivant la loi d'action de masse $p_0 \approx 10^4$ cm^{-3}. Supposons maintenant que près de la surface est créée, sous l'effet d'un rayonnement lumineux, une concentration excédentaire des porteurs $\Delta n = \Delta p = 10^{14}$ cm^{-3}. Si la durée de vie des porteurs hors d'équilibre est égale à 1 μs, la longueur de diffusion est $L \approx 0{,}0035$ cm. Alors le gradient de concentration pour $x = 0$ est égal à $\Delta n/L \approx 3 \cdot 10^{16}$ cm^{-4} et, suivant (8.40), l'intensité du champ « incorporé » vaut $\sim$ 0,05 V/cm. Dans un tel champ, la composante de dérive de la densité de courant des porteurs minoritaires — des trous — est $i_{p\ der} \approx 3{,}5 \cdot 10^{-4}$ A/cm^2 [les mobilités u_n et u_p sont prises égales respectivement à 1300 et 500 cm^2/(V·s)] et la composante de diffusion $i_{p\ diff} \approx 6 \cdot 10^{-2}$ A/cm^2.

On peut en conclure que les porteurs minoritaires ne « ressentent » pas en fait le champ électrique créé et que leur mouvement s'effectue uniquement par diffusion. Il n'en est plus de même pour les porteurs majoritaires. Comme ils sont beaucoup plus nombreux, leurs courants de dérive sont beaucoup plus intenses que les courants de dérive des porteurs minoritaires.

Regardons maintenant comment est justifiée l'hypothèse que les concentrations des électrons et des trous hors d'équilibre sont pratiquement égales, c'est-à-dire que $\Delta n = \Delta p$ par suite de l'égalisation de leurs courants par le champ « incorporé ». Pour que le champ $\mathscr{E}$ existe, il faut que la compensation des électrons par les trous ne soit pas complète. Les charges et les champs sont liés par l'*équation de poisson* :

$$\frac{d^2V}{dx^2} = -\frac{d\mathscr{E}}{dx} = -\frac{\rho(x)}{\varepsilon\varepsilon_0}, \tag{8.46}$$

où $\rho(x)$ est la densité de la charge d'espace ; V le potentiel du champ électrique et $\varepsilon\varepsilon_0$ pour le silicium est de 10^{-12} F/cm. Etant donné qu'en l'absence de porteurs hors d'équilibre dans un semiconducteur homogène est partout satisfaite la condition d'électroneutralité, une charge d'espace prend naissance si $\Delta n \neq \Delta p$:

$$\rho(x) = e(\Delta p - \Delta n). \tag{8.47}$$

Dans l'exemple considéré $|d\mathscr{E}/dx|$ est maximal pour $x = 0$ et est égal à 10 V/cm^2. De (8.46) et (8.47) on obtient alors que $|(\Delta p - \Delta n)| \approx 10^8$ cm^{-3}. Cette valeur est réellement beaucoup plus petite que les concentrations excédentaires elles-mêmes égales à $\sim 10^{14}$ cm^{-3}.

Examinons enfin la rapidité avec laquelle est compensée la charge des porteurs d'un même signe injectés dans le semiconducteur. Négligeons la recombinaison des porteurs au cours du processus examiné. Alors, les équations de continuité dans le cas unidimension-

nel prennent la forme

$$\begin{cases} \dfrac{\partial(\Delta p)}{\partial t} = -\dfrac{1}{e}\dfrac{\partial i_p}{\partial x}, \\ \dfrac{\partial(\Delta n)}{\partial t} = \dfrac{1}{e}\dfrac{\partial i_n}{\partial x}. \end{cases} \tag{8.48}$$

La vitesse de disparition de la charge d'espace dont la densité $\rho(x) = e(\Delta p - \Delta n)$ est obtenue en retranchant de la première équation (8.48) la seconde:

$$\frac{\partial}{\partial t}[e(\Delta p - \Delta n)] = \frac{\partial \rho}{\partial t} = -\frac{\partial}{\partial x}(i_p + i_n). \tag{8.49}$$

Tant que le champ électrique engendré par les charges d'espace est relativement intense, on peut considérer que les courants circulent dans le semiconducteur principalement par dérive. Alors

$$i = i_p + i_n = \sigma\mathscr{E}, \tag{8.50}$$

où σ est la conductivité électrique du semiconducteur. Si le niveau d'injection est faible, c'est-à-dire si $(\Delta n + \Delta p) \ll (n_0 + p_0)$, on peut considérer que $\sigma \simeq \sigma_0$ et alors

$$\frac{\partial i}{\partial x} = \sigma_0 \frac{\partial \mathscr{E}}{\partial x}. \tag{8.51}$$

Exprimons $\partial\mathscr{E}/\partial x$ en fonction de la densité de la charge d'espace suivant l'équation de Poisson (8.46), il vient

$$\frac{\partial \rho}{\partial t} = -\frac{\sigma_0}{\varepsilon\varepsilon_0}\rho. \tag{8.52}$$

La solution de cette équation est de la forme

$$\rho(t) = \rho(0)\, e^{-\sigma_0 t/(\varepsilon\varepsilon_0)} = \rho(0)\, e^{-t/\tau_M}. \tag{8.53}$$

Ici, $\rho(0)$ désigne la densité initiale de la charge d'espace et

$$\tau_M = \varepsilon\varepsilon_0/\sigma_0 \tag{8.54}$$

est un *temps de relaxation* dit *de Maxwell.*

Pour le silicium, même de faible conductivité électrique $\sigma_0 = 1\ \text{ohm}^{-1}\cdot\text{cm}^{-1}$, le temps de relaxation $\tau_M \approx 10^{-12}$ s. On voit que les temps de relaxation des charges d'espace dans les semiconducteurs sont très petits. A propos, cela montre qu'on a eu raison en négligeant plus haut la recombinaison des porteurs hors d'équilibre durant le temps de relaxation. Il convient de noter que bien que le temps de relaxation de Maxwell ne dépende pas du type de porteurs injectés (majoritaires ou minoritaires) qui donnent lieu à la formation d'une charge d'espace, cette relaxation se produit toujours

aux frais des courants de porteurs majoritaires. C'est pourquoi, lorsque les porteurs injectés sont minoritaires, les porteurs majoritaires viennent les compenser. Si on injecte des porteurs majoritaires, il se produit une résorption, une redistribution de leur charge. Dans le premier cas, au bout du temps τ_M la concentration augmentée des porteurs minoritaires est conservée et des porteurs majoritaires venant de la profondeur du semiconducteur ou du circuit extérieur s'approchent des porteurs minoritaires de sorte que $\Delta n = \Delta p$ tandis que dans le second cas les porteurs hors d'équilibre se répartissent sur tout l'échantillon ou s'en vont dans le circuit extérieur et $\Delta n = \Delta p = 0$.

CHAPITRE 9

PHÉNOMÈNES OPTIQUES ET PLASMIQUES DANS LES SEMICONDUCTEURS

Les effets optiques que présentent les semiconducteurs sont utilisés dans les divers photorécepteurs, les convertisseurs de l'énergie solaire en énergie électrique, les sources de radiations à semiconducteurs et dans certains autres appareils opto-électroniques. Les principaux types de ces appareils seront étudiés au cours des chapitres qui suivent. Quant au chapitre actuel, il se propose d'examiner la nature des processus optiques les plus importants qui sont l'absorption et l'émission de la lumière ainsi que le principe de fonctionnement de la photorésistance qui est le photorécepteur le plus simple (mais très important en pratique).
La deuxième partie de ce chapitre est consacrée à l'étude des effets acousto-électriques et des effets de réchauffage du gaz électronique dans des champs électriques intenses qui sont eux aussi largement utilisés dans l'électronique moderne.

§ 9.1. Absorption de la lumière

Coefficient d'absorption de la lumière. Dirigeons un faisceau de lumière de puissance W_0 (fig. 9.1) sur un semiconducteur homogène. A mesure que la lumière pénètre dans le semiconducteur, elle s'absorbe progressivement et sa puissance diminue. Découpons, à une profondeur x à partir de la surface, une mince couche semiconductrice dx. La puissance de l'énergie lumineuse dW absorbée par la couche dx est proportionnelle à la puissance W de la lumière tombant sur cette couche et à son épaisseur :

$$dW = -\alpha W\, dx. \tag{9.1}$$

Le signe moins indique que la puissance diminue. Le coefficient de proportionnalité α est appelé *coefficient d'absorption*. Pour $dx = 1$ on a $\alpha = -dW/W$. *Le coefficient d'absorption est donc numériquement égal à la variation relative de puissance de la lumière sur un chemin unité parcouru dans le milieu absorbant.* La dimension du coefficient d'absorption est l'inverse d'une longueur (m^{-1}, cm^{-1}).

En intégrant (9.1) compte tenu de la réflexion sur la surface, on obtient

$$W = W_0 (1 - r)e^{-\alpha x}, \tag{9.2}$$

où r est le coefficient de réflexion de la lumière. Dans des milieux non homogènes il se produit non seulement l'absorption de la lumière, mais également la diffusion de la lumière sur les pores, les inclusions, etc., mais de tels milieux ne seront pas considérés dans la suite de cet ouvrage.

Proposons-nous maintenant d'examiner la nature des divers mécanismes de l'absorption de la lumière dans les semiconducteurs.

Absorption propre. Lors de l'*absorption propre* l'énergie de la lumière pénétrant dans le semiconducteur est dépensée pour l'excitation des électrons de la bande de valence dans la bande de conduction (fig. 9.2). En raison de la loi de la conservation de l'énergie,

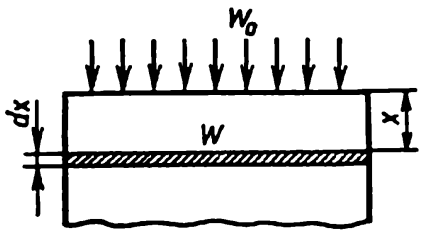

Fig. 9.1. Explication de la loi d'absorption de la lumière dans la substance.

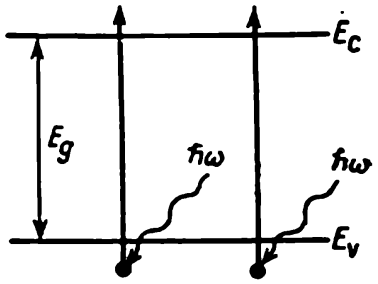

Fig.9 .2. Schéma de l'absorption propre de la lumière dans les semiconducteurs.

une telle absorption ne peut se produire que dans le cas où l'énergie $\hbar\omega$ des quanta lumineux n'est pas inférieure à la largeur E_g de la bande interdite:

$$\hbar\omega \geqslant E_g. \tag{9.3}$$

A partir de cette condition on peut déterminer la longueur d'onde maximale λ_{max} de l'absorption propre:

$$\lambda_{max} = 2\pi c/\omega = 2\pi c\hbar/E_g, \tag{9.4}$$

où c est la vitesse de la lumière. Pour le silicium par exemple, dont $E_g \approx 1{,}1$ eV, on a $\lambda_{max} \approx 1{,}13$ μm.

Le traitement du processus d'absorption de la lumière selon la théorie quantique montre qu'en plus de la loi de la conservation de l'énergie c'est la loi de la conservation de l'impulsion qui doit être observée:

$$\mathbf{p}_n = \mathbf{p}_p + \mathbf{p}_{phot}. \tag{9.5}$$

Ici $\mathbf{p}_n = \hbar\mathbf{k}_n$ est l'impulsion de l'électron transféré dans la bande de conduction; $\mathbf{p}_p$ l'impulsion du trou qui apparaît dans la bande de valence; $\mathbf{p}_{phot}$ l'impulsion du photon (quantum de lumière) qui a produit le passage de l'électron.

Dans les limites de la première zone de Brillouin, les projections de l'impulsion de l'électron sur les axes cristallographiques sont comprises entre $-\hbar\pi/a$ et $+\hbar\pi/a$, où a est le paramètre du réseau ($a \approx 3\cdot10^{-8}$ cm, $\hbar\pi/a \approx 10^8\hbar$). L'impulsion du photon est égale à $2\pi\hbar/\lambda$ et, pour $\lambda \approx 5\cdot10^{-5}$ cm, elle vaut $\sim 10^5\,\hbar$, c'est-à-dire est d'environ trois ordres de grandeur inférieure à l'impulsion de l'élec-

tron. On peut donc considérer que dans les transitions optiques l'impulsion de l'électron reste pratiquement inchangée:

$$\hbar \mathbf{k}_n \approx \hbar \mathbf{k}_p. \tag{9.6}$$

Dans le diagramme des énergies de telles transitions (dites *directes*) sont représentées par des flèches verticales (*1*, *2* sur la figure 9.3, *a*).

Le calcul théorique du coefficient d'absorption propre pour des transitions directes dans des semiconducteurs avec des extrémums

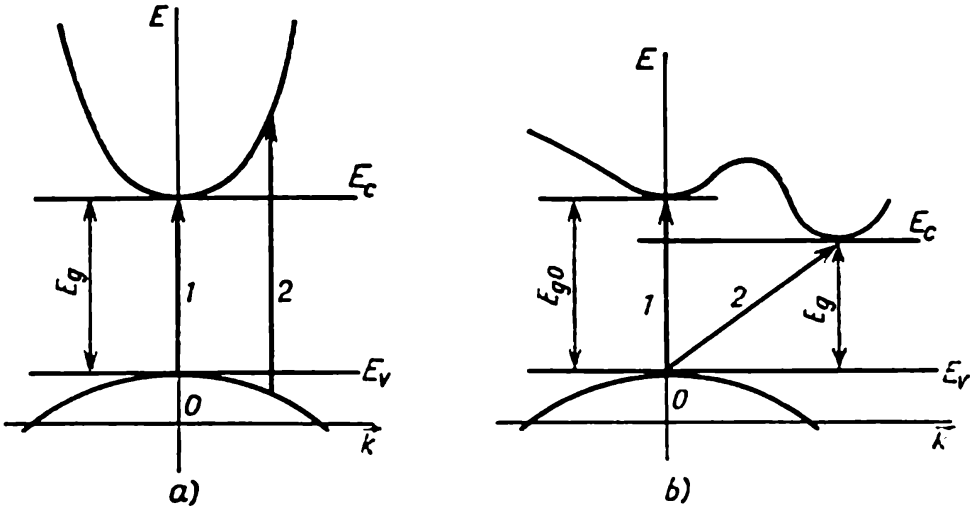

Fig. 9.3. Transitions directes (*a*) et indirectes (*b*) lors de l'absorption propre de la lumière.

des bandes situés pour une seule et même valeur de **k** (fig. 9.3, *a*) conduit à l'expression suivante:

$$\alpha_{pr} = \frac{e^2 [2m_n m_p/(m_n + m_p)]^{3/2}}{4\pi \bar{n} c \hbar^2 \varepsilon_0 m_c} (\hbar\omega - E_g)^{1/2}, \tag{9.7}$$

où $\bar{n}$ est l'indice de réfraction du semiconducteur. En posant $\bar{n} = 4$ et les masses effectives des électrons m_n et des trous m_p égales à la masse d'un électron libre et en exprimant $\hbar\omega$ et E_g en eV et α_{pr} en cm^{-1}, on obtient

$$\alpha_{pr} \approx 2{,}7 \cdot 10^5 (\hbar\omega - E_g)^{1/2}.$$

En effet, dans le domaine de l'absorption propre le coefficient α_{pr} atteint des valeurs comprises dans les limites de $\sim 10^4$ à 10^5 cm^{-1}, ce qui signifie que la lumière est déjà absorbée à une profondeur de pénétration de $\sim 1{,}0$ à $0{,}1$ µm.

Si le niveau inférieur de la bande de conduction E_c se situe pour une autre valeur de **k** que le sommet de la bande de valence E_v (fig. 9.3, *b*), comme c'est le cas pour le germanium et le silicium, la

distance E_{g0}, suivant la verticale, entre les bandes est plus grande que la largeur de la bande interdite $E_g = E_c - E_v$. Dans ce cas, les transitions optiques directes ne peuvent être excitées que par des quanta lumineuses d'énergie supérieure à E_{g0}:

$$\hbar\omega \geqslant E_{g0}. \tag{9.8}$$

La grandeur E_{g0} est appelée *largeur optique* de la bande interdite, elle est égale à la distance minimale entre les bandes suivant la verticale.

Dans de tels semiconducteurs en plus des transitions directes peuvent se produire aussi des *transitions indirectes* montrées sur la figure 9.3, *b* par une flèche inclinée *2*. Elles se font avec la participation d'une troisième quasi-particule, du phonon. Dans ce cas, les lois de la conservation de l'énergie et de l'impulsion prennent la forme suivante:

$$E_n = E_p + \hbar\omega \pm E_{phon}, \tag{9.9}$$

$$\mathbf{p}_n = \mathbf{p}_p + \mathbf{p}_{phot} \pm \mathbf{p}_{phon}. \tag{9.10}$$

Ici, le signe plus se rapporte à des processus se déroulant avec l'absorption d'un phonon et le signe moins à des processus se produisant avec l'émission d'un phonon. Comme l'énergie des phonons dans les semiconducteurs ne dépasse pas quelques centièmes d'électron-volt et $\hbar\omega \approx 1$ eV, l'énergie E_{phon} figurant dans l'expression (9.9) peut être négligée devant $\hbar\omega$. Quant à l'impulsion du phonon $\hbar\mathbf{k}_{phon}$, elle se situe dans les limites de la première zone de Brillouin de même que l'impulsion de l'électron. Aussi, dans les transitions avec la participation des phonons l'impulsion de l'électron peut-elle varier dans de larges limites, ce qu'on exprime graphiquement par le tracé des flèches inclinées qui caractérisent de telles transitions (fig. 9.3, *b*).

La probabilité de processus se déroulant avec la participation de trois particules étant beaucoup plus faible que celle des processus à deux particules, le coefficient d'absorption est beaucoup plus bas dans le domaine des transitions indirectes que dans celui des transitions directes. Lorsque la température baisse, les processus avec l'absorption d'un phonon deviennent plus rares et le coefficient d'absorption pour les transitions indirectes diminue.

Le bord d'absorption propre dans les semiconducteurs se déplace un peu sous l'effet de la pression qui provoque une variation de la constante de réseau et, avec elle, de la structure énergétique du semiconducteur. Lors de la compression triaxiale la variation de E_g peut être positive ou négative. Aussi, le bord d'absorption propre peut-il se déplacer aussi bien vers la partie « ondes longues » que vers la partie « ondes courtes » du spectre.

Le même effet détermine aussi la variation de la structure de bande du semiconducteur dans le cas d'une variation de température qui conduit au déplacement du bord d'absorption propre.

L'absorption au voisinage du bord du domaine propre subit aussi une variation sous l'effet d'un champ électrique extérieur. Ce phénomène, qui a reçu le nom d'*effet Franz-Keldych*, s'explique par l'inclinaison que les bandes d'énergie présentent dans un champ électrique extérieur $\mathscr{E}$ (fig. 9.4) et par pénétration, par effet tunnel, des électrons à la profondeur $x_3 - x_1$ depuis le point x_1 vers la droite et à la distance $x_2 - x_3$ depuis le point x_2 vers la gauche. Comme le

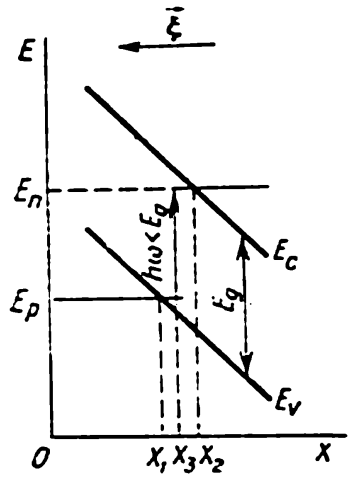

Fig. 9.4. Effet Franz-Keldych.

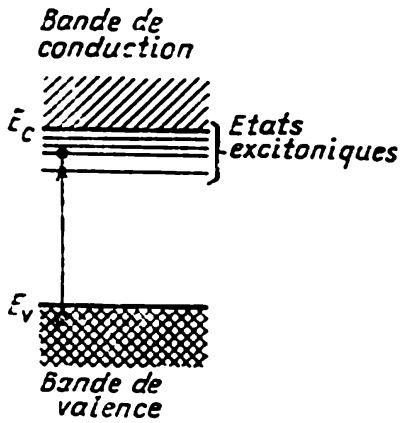

Fig. 9.5. Absorption excitonique de la lumière.

montre la figure 9.4, l'électron situé au point x_3 peut passer de l'état E_p dans l'état E_n en absorbant un quantum de lumière $\hbar\omega = E_n - E_p < E_g$.

Considérons l'influence du dopage du semiconducteur sur l'absorption propre. Tant que le niveau de dopage n'est pas trop élevé et le semiconducteur reste non dégénéré, le dopage est pratiquement sans effet sur le spectre d'absorption propre. Cela s'explique par le fait que dans les semiconducteurs non dégénérés le degré d'occupation par les électrons des états dans la bande de conduction est très faible de sorte qu'ils ne contrarient pas les passages des électrons de la bande de valence. D'un autre côté, dans les semiconducteurs non dégénérés même de type p, le degré d'occupation des états dans la bande de valence est voisin de l'unité et la probabilité de transitions optiques à partir de ces états ne dépend pas du degré de dopage.

La situation est différente dans les semiconducteurs dégénérés. Une faible dégénérescence conduit à une diminution du coefficient d'absorption à des fréquences voisines du bord d'absorption propre. Pour ce qui est de la dégénérescence forte, elle provoque le déplacement du bord d'absorption vers des longueurs d'ondes plus courtes. Cet effet porte le nom de *déplacement de Burstein*. Il se manifeste nettement dans les semiconducteurs à faible densité des états près du

niveau inférieur de la bande de conduction (ou près du sommet de la bande de valence) dans lesquels une forte dégénérescence est obtenue pour des niveaux de dopage relativement bas. C'est ainsi que dans InSb le dopage avec des impuretés donatrices (concentration de $5 \cdot 10^{24}$ m^{-3}) a pour résultat le déplacement de la frontière ondes longues de l'absorption propre de 7,1 à 3,5 μm. Mais, dans de nombreux cas, le déplacement de Burstein est masqué par un autre effet dû à un fort dopage — la variation de la densité d'états près des bords des bandes d'énergie. Cette variation se produit par suite de l'étalement des niveaux d'impureté qui se transforment en une bande d'impureté qui se confond avec la bande de conduction ou avec la bande de valence.

En outre, au voisinage des atomes d'impureté le réseau cristallin se trouve comprimé ou dilaté suivant le rapport entre les rayons des atomes principaux et les atomes d'impureté. Comme la largeur de la bande interdite E_g varie avec la pression, ceci doit conduire à la fluctuation de $_g$ le long du cristal et à une variation du spectre d'absorption.

Enfin, un champ électrique régnant au voisinage des atomes ionisés d'impureté doit conduire, en vertu de l'effet Franz-Keldych, au déplacement du bord d'absorption dans la région des ondes longues.

Excitons. Dans le cas de l'absorption propre de la lumière, les électrons de la bande de valence sont transférés dans la bande de conduction et deviennent libres. Toutefois le processus peut se dérouler d'une autre façon lorsque l'électron excité ne rompt pas sa liaison avec le trou qui apparaît dans la bande de valence mais forme avec lui un système lié unique. Un tel système a été étudié pour la première fois par Y. Frenkel auquel il a donné le nom d'*exciton*. L'exciton est semblable à l'atome d'hydrogène: dans les deux cas un électron se déplace autour d'une charge positive unité et le spectre énergétique est discret (fig. 9.5). Les niveaux d'énergie de l'exciton se situent près du niveau inférieur de la bande de conduction.

Comme les excitons sont des systèmes électriquement neutres, leur apparition dans le semiconducteur ne conduit pas à l'apparition de porteurs de charge additionnels de sorte que l'absorption de la lumière ne s'accompagne pas d'une augmentation de la conductivité du semiconducteur. Mais lors des collisions avec les phonons, les atomes d'impureté et d'autres défauts du réseau cristallin les excitons se recombinent ou se désintègrent. Dans le premier cas l'électron excité passe dans la bande de valence et l'énergie d'excitation est transmise au réseau ou émise sous forme de quanta de lumière; dans le second cas il y a formation d'une paire de porteurs — d'un électron et d'un trou — ce qui a pour résultat une augmentation de la conductibilité électrique du semiconducteur.

Absorption de la lumière par des porteurs de charges libres. La lumière peut provoquer des transitions des porteurs de charges libres

d'un niveau à un autre de la bande (fig. 9.6, *a*). Etant donné pue dans de telles transitions l'impulsion du porteur doit subir une variation notable, elles ne peuvent se produire qu'avec la participation d'une troisième particule.

Selon la théorie classique, l'absorption de la lumière par des porteurs libres se produit comme suit : les porteurs de charges sont accélérés dans le champ électrique de l'onde lumineuse et, en se diffusant sur les défauts du réseau cristallin, leur transmettent leur énergie. Autrement dit l'énergie de l'onde lumineuse se transforme en chaleur par effet Joule-Lenz. La formule classique pour le coefficient d'absorption de la lumière par les porteurs libres se présente sous la forme suivante :

$$\alpha_{lib} = e^3 n \lambda^2 / (\bar{n} c^3 u_0 m_n 4\pi^2 \varepsilon_0), \quad (9.11)$$

où c est la vitesse de la lumière dans le vide ; $\bar{n}$ l'indice de réfraction du semiconducteur ; u_0 la mobilité des porteurs de charges en courant continu ; ε_0 la permittivité du vide.

L'expérience confirme que α_{lib} est proportionnel à la concentration n de porteurs libres tant que la mobilité u_0 ne commence à varier avec l'augmentation du dopage ; pour certains semiconducteurs la dépendance $\alpha_{lib} \sim \lambda^2$ est valable.

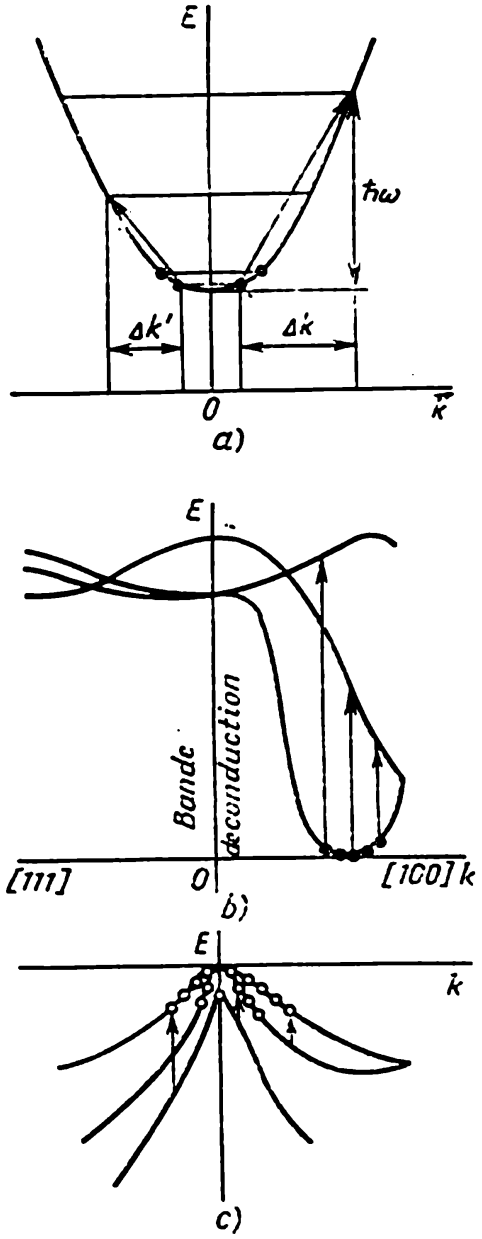

Fig. 9.6. Absorption de la lumière par les porteurs de charges libres lors des transitions à l'intérieur d'une sous-bande (*a*), entre les sous-bandes de la bande de conduction (*b*) et la bande de valence (*c*) du silicium.

Transitions entre sous-bandes. La structure de la bande de conduction et de la bande de valence d'un semiconducteur réel est plus complexe que celle décrite plus haut. Ces bandes sont représentées par une superposition de plusieurs bandes (que l'on appelle plus souvent *sous-bandes*) de sorte que pour chaque bande il existe non pas une seule mais plusieurs courbes de dispersion

qui se croisent parfois. Les porteurs de charges libres — les électrons dans un semiconducteur de type n et les trous dans un semiconducteur de type p — peuvent passer entre ces sous-bandes en absorbant la lumière comme le montrent les figures 9.6, *b* et *c*. Dans un tel cas, il apparaît dans le spectre d'absorption des bandes supplémentaires qui se superposent au spectre (9.11). La forme et l'intensité de ces bandes dépendent du degré de dopage du semiconducteur. La figure 9.7 montre à titre d'exemple les spectres d'absorption pour le GaAs — n avec $N_d = 1{,}3 \cdot 10^{17}$ et $5{,}4 \cdot 10^{18}$ cm^{-3}.

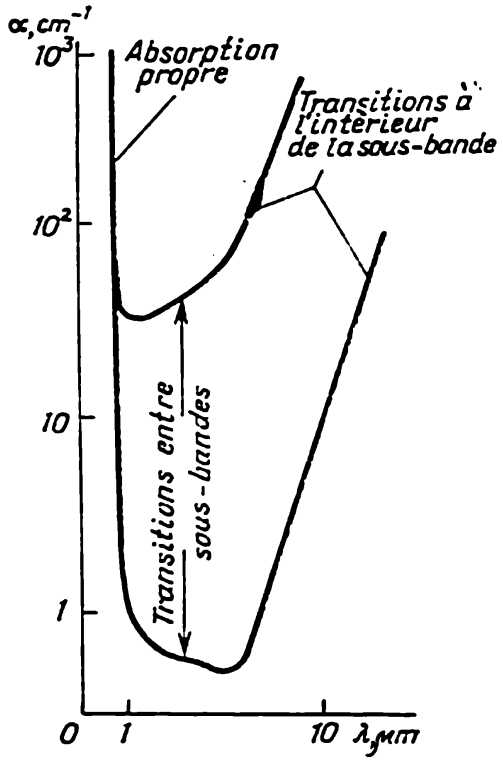

Fig. 9.7. Spectre d'absorption de la lumière par l'arséniure de gallium n.

Absorption due aux impuretés. Dans les semiconducteurs dopés les électrons peuvent passer sous l'effet de la lumière des niveaux d'impureté dans la bande de conduction (fig. 9.8, *a*) et de la bande de valence aux niveaux d'impureté situés dans la bande interdite (fig. 9.8, *b*). Une telle absorption de la lumière est dite *due aux impuretés*. La frontière de cette absorption est déplacée dans le domaine des ondes longues d'autant plus fortement que l'énergie de la transition correspondante est plus petite.

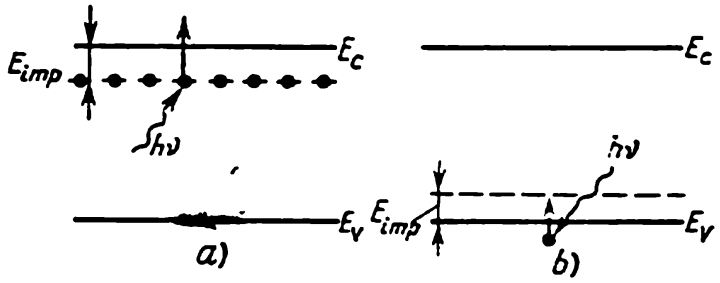

Fig. 9.8. Absorption de la lumière due aux impuretés.

Il faut pourtant ne pas perdre de vue que si les atomes d'impureté sont déjà ionisés, l'absorption due aux impuretés n'aura pas lieu. Comme la température d'appauvrissement de l'impureté tombe lorsque son énergie d'ionisation diminue, pour pouvoir observer

l'absorption ondes longues due aux impuretés, le semiconducteur doit être refroidi à une température suffisamment basse.

C'est ainsi que le spectre d'absorption due aux impuretés d'un germanium dopé à l'or (l'énergie d'ionisation de l'impureté $E_i = 0{,}08$ eV, la frontière d'absorption $\lambda_{max} \approx 9$ μm) s'observe à la température de l'azote liquide (77 K) alors que pour un germanium dopé à l'antimoine ($E_i = 0{,}01$ eV, $\lambda_{max} = 135$ μm) l'absorption due aux impuretés ne peut être observée qu'à des températures de l'hélium liquide (4 K).

Le coefficient d'absorption due aux impuretés α_{imp} pour $\lambda < \lambda_{max}$ dépend de λ et est proportionnel à la concentration de l'impureté N_{imp}. Aussi, l'absorption due aux impuretés peut-elle être commodément caractérisée par la *section différentielle d'absorption* σ_{imp}:

$$\sigma_{imp} = \alpha_{imp}(\lambda)/N_{imp}. \tag{9.12}$$

Cette section atteint sa valeur maximale près du bord de l'absorption due aux impuretés. Au point de maximum, $\sigma_{imp} \approx 10^{-16}$ à 10^{-15} cm^2 de sorte que $\alpha_{imp} \approx 1$ à 10 cm^{-1} pour $N_{imp} = 10^{16}$ cm^{-3}. Parfois on n'arrive pas à obtenir de grandes valeurs de α_{imp} du fait des petites limites de solubilité des impuretés dans les semiconducteurs. Par exemple la limite de solubilité de l'or dans le germanium est justement égale à $\sim 10^{16}$ cm^{-3}.

§ 9.2. Emission d'un rayonnement lumineux par les semiconducteurs

La recombinaison des paires électron-trou dans les semiconducteurs peut s'accompagner d'une émission de rayonnement lumineux (voir § 8.2). Les processus d'émission sont des inverses des processus d'absorption et, à l'état d'équilibre thermique, le nombre de quanta de lumière émis par le semiconducteur est exactement égal au nombre de quanta qu'il absorbe. L'apparition dans le semiconducteur de porteurs de charges hors d'équilibre, en excès, fait prédominer les processus d'émission sur les processus d'absorption et le semiconducteur commence à s'illuminer.

L'émission de la lumière peut se produire aussi bien lors de la recombinaison bande à bande que lors de la première ou de la deuxième (ou de la première et de la seconde) étape de la recombinaison par les pièges (flèches *2* et *3* de la figure 8.2). Toutefois, les transitions radiatives sont le plus souvent celles bande à bande ou celles entre un niveau donneur peu profond et la bande de valence ou entre la bande de conduction et un niveau accepteur peu profond (fig. 9.9, *a*, *b* et *c* respectivement). La probabilité de telles transitions est la plus forte dans des matériaux à bandes directes tels que l'arséniure de gallium, bien que la lueur de recombinaison peut s'observer aussi dans le germanium et le silicium.

Comme il a été montré au § 9.1, un électron libre et un trou libre peuvent former un exciton. Si, après cela, l'électron et le trou se recombinent dans l'exciton et il y a annihilation de celui-ci, un tel processus s'accompagne en règle générale d'une émission d'un quantum de lumière (fig. 9.9, *d*). Les excitons peuvent se déplacer librement dans le cristal ou être liés à des défauts de toute nature de la structure cristalline. Ces défauts peuvent être représentés par les atomes d'impureté dopeuse, c'est-à-dire les niveaux donneurs et accepteurs

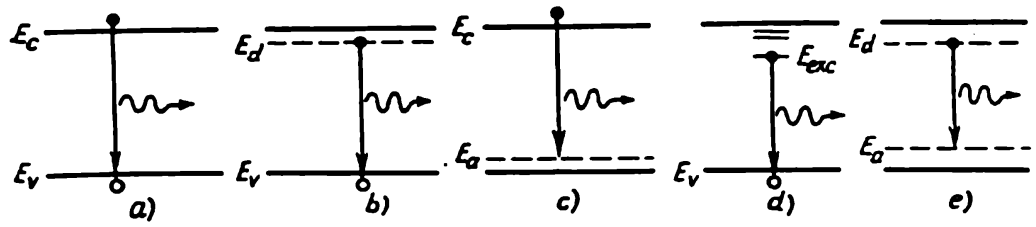

Fig. 9.9. Schémas des transitions radiatives dans les semiconducteurs:
a — bande à bande ; *b* — transition d'un niveau donneur peu profond dans la bande de valence ; *c* — transition de la bande de conduction à un niveau accepteur ; *d* — annihilation d'un exciton ; *e* — transition d'un donneur à un accepteur.

peu profonds, ainsi que par les atomes d'impureté neutre, par exemple les atomes d'azote qui se substituent aux atomes de phosphore dans le phosphure de gallium. L'énergie d'un exciton est plus faible à l'état lié qu'à l'état libre de sorte que les excitons liés sont moins détruits par l'agitation thermique et la recombinaison par les excitons liés peut augmenter brusquement la part des actes de recombinaison se produisant avec émission de la lumière surtout dans les semiconducteurs à bandes non directes.

Enfin les transitions des électrons d'un niveau donneur à un niveau accepteur peuvent elles aussi être radiatives (fig. 9.9, *e*). Mais pour que de telles transitions soient possibles, le semiconducteur doit être fortement dopé afin que les atomes donneurs et les atomes accepteurs se situent souvent à une distance suffisamment petite les uns des autres.

§ 9.3. Photoconductivité des semiconducteurs

Les effets d'absorption de la lumière dans les semiconducteurs trouvent des applications pratiques pour la réalisation de modulateurs de flux lumineux, de photorécepteurs et de convertisseurs de l'énergie lumineuse en énergie électrique.

Pour la modulation des flux lumineux on peut utiliser la variation du coefficient d'absorption près du bord de la bande propre en fonction de la pression, du champ électrique extérieur et de la concentration de porteurs de charges libres ou la variation de l'absorption de la lumière par les porteurs libres en fonction de leur concentration.

L'absorption propre et l'absorption due aux impuretés provoquent l'apparition de porteurs de charges libres en excès dont la présence fait croître la conductivité du semiconducteur. Le processus de libération interne des électrons sous l'action de la lumière porte le nom d'*effet photoélectrique interne*. L'accroissement de conductivité électrique acquise par le semiconducteur par suite de son irradiation par la lumière est appelé *photoconductivité*. Quant à la conduction principale, due à l'excitation thermique des porteurs de charges libres, elle est appelée *conduction d'obscurité*. Les appareils permettant d'enregistrer le rayonnement lumineux à l'aide du phénomène de photoconduction portent le nom de *photorésistances* ou de *cellules photoconductives*.

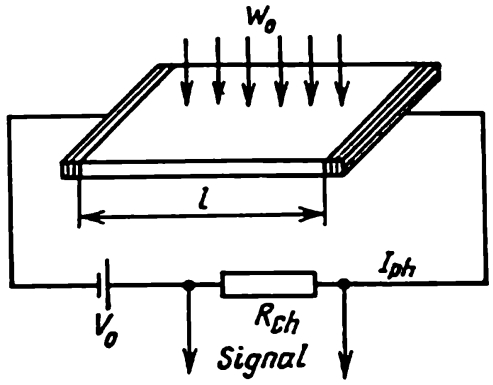

Fig. 9.10. Schéma de constitution et de branchement d'une photorésistance.

La figure 9.10 montre la constitution d'une photorésistance et le schéma de son branchement. L'élément sensible de la photorésistance est constitué par un barreau ou une pellicule de semiconducteur monocristallin ou polycristallin comportant deux contacts ohmiques. Il est relié à une source de polarisation V_0 via une résistance de charge R_{ch}. L'élément sensible doit être suffisamment épais pour pouvoir absorber pratiquement toute la lumière W_0 $(1 - r)$ ayant traversé la surface éclairée (W_0 est la puissance de la lumière incidente; r le coefficient de réflexion de la surface). Cette condition est facile à satisfaire dans le cas de photorésistances intrinsèques et difficile à accomplir pour les photorésistances extrinsèques. Si cette condition est réalisée, le nombre de porteurs (ou de paires de porteurs dans le cas de l'absorption propre) produits par la lumière en unité de temps dans l'élément sensible pour $\lambda < \lambda_{max}$ est égal à

$$G = W_0 (1 - r) \eta/(\hbar\omega), \tag{9.13}$$

où η est le *rendement quantique* de l'effet photoélectrique interne, égal au nombre de porteurs (de paires de porteurs) produit en moyenne par chaque photon absorbé. Il peut être supérieur à 1, si l'absorption d'un photon de haute énergie fait naître deux (ou un plus grand nombre) de paires électron-trou et inférieur à 1, si une partie de photons est absorbée par des porteurs de charges libres.

Sous l'action de la tension V_0, appliquée à la photorésistance, les porteurs de charges créés par la lumière sont animés d'un mouvement de dérive en faisant circuler dans le circuit un courant que

l'on appelle *courant photoélectrique* I_{ph}. Ce courant est facile à déterminer en partant des considérations suivantes. Pendant la durée de sa vie τ chacun des porteurs de charge traverse la résistance τ/t_{tr} fois, où t_{tr} est le temps de transit ou plus exactement le temps mis par le porteur pour traverser la résistance par dérive *). Il est égal à la longueur l de l'élément sensible de la résistance divisée par la vitesse de dérive v_{der}:

$$t_{tr} = l/v_{der} = l/(u\mathscr{E}) = l^2/(uV), \tag{9.14}$$

où u est la mobilité des porteurs. L'intensité de courant photoélectrique I_{ph} est égale au produit du nombre G de porteurs produits chaque seconde dans le semiconducteur sous l'effet de la lumière par le rapport τ/t_{tr} et par la charge e de l'électron:

$$I_{ph} = eG\tau/t_{tr} = eG\tau uV/l^2 = e\Delta NuV/l^2, \tag{9.15}$$

où $\Delta N = G\tau$ est le nombre de porteurs en excès dans la photorésistance. Si la lumière produit une paire de porteurs, au lieu de u il convient d'écrire la somme des mobilités $u_n + u_p$.

En portant dans (9.15) la valeur de G tirée de (9.13), on obtient

$$I_{ph} = \frac{W_0(1-r)e\eta\tau uV}{\hbar\omega l^2} = \frac{W_0(1-r)e\eta\tau\lambda uV}{hcl^2}. \tag{9.16}$$

Le rapport

$$\frac{I_{ph}}{W_0} = \frac{e(1-r)}{hcl^2}\eta\tau u\lambda V \tag{9.17}$$

caractérise la sensibilité de la photorésistance qui est proportionnelle à la longueur d'onde de la lumière incidente (jusqu'à λ_{max}), à la tension appliquée V, à la durée de vie τ des porteurs et à leur mobilité u et inversement proportionnelle au carré de la longueur de l'élément sensible de la photorésistance.

La durée de vie τ des porteurs détermine non seulement la sensibilité mais aussi l'inertie de la photorésistance: plus longue est la durée de vie τ, plus grande est la sensibilité, mais plus grande aussi est l'inertie de l'appareil. En effet, considérons par exemple le caractère de décroissance de la photoconductivité de la photorésistance après le débranchement de la source de lumière (fig. 9.11). Par suite du processus de recombinaison la concentration des porteurs en excès décroît suivant la loi [v. (8.10)]

$$\Delta n = \Delta n_0 e^{-(t-t_0)/\tau},$$

*) A proprement parler, lorsqu'un porteur de charge sort de la résistance, un autre porteur y entre par le deuxième contact, mais cela ne change pas le fond de l'affaire.

où t_0 est l'instant de débranchement de la lumière. La photoconductivité du semiconducteur décroît suivant une même loi (de e fois durant le temps τ; courbe BC de la figure 9.11):

$$\sigma(t) = \sigma_{ph}(t_0)\, e^{\frac{-t-t_0}{\tau}}, \tag{9.18}$$

où $\sigma_{ph}(t_0)$ est la photoconductivité à l'instant de débranchement de la lumière.

La formule (9.18) montre que plus grande est la durée de vie des porteurs en excès, plus lente est la décroissance de la photoconducti-

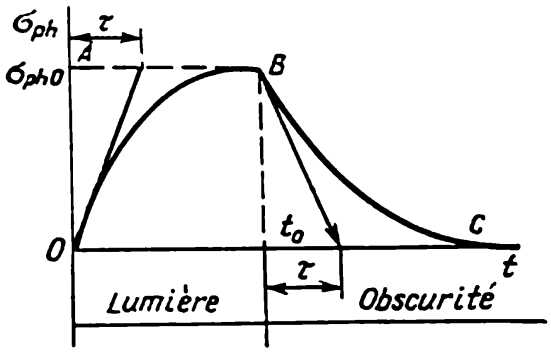

Fig. 9.11. Courbe de croissance de la photoconduction lors de l'irradiation du semiconducteur par la lumière et de décroissance de la photoconduction après suppression de l'irradiation.

vité et donc plus grande est l'inertie du photorécepteur de rayonnement.

Il est facile de montrer que la tangente menée en un point t_0 à la courbe de décroissance de la photoconduction $\sigma_{ph}(t)$ coupe sur l'axe des temps un tronçon numériquement égal à τ — la durée de vie des porteurs en excès. Cette circonstance est souvent utilisée pour une détermination expérimentale de τ.

La figure 9.11 montre aussi le caractère de croissance de la photoconductivité d'un semiconducteur après l'application d'une impulsion lumineuse (courbe OB) décrite par l'équation suivante:

$$\sigma_{ph} = \sigma_{ph\,0}(1 - e^{-t/\tau}), \tag{9.19}$$

où $\sigma_{ph\,0}$ est la *photoconductivité stationnaire* (en régime permanent) dans les conditions d'éclairement constant de la photorésistance. Dans ce cas lui aussi, la tangente menée à la courbe $\sigma_{ph}(t)$ à l'origine des coordonnées coupe sur la droite AB un tronçon numériquement égal à τ.

Il convient d'indiquer que dans les photorésistances polycristallines la sensibilité est souvent déterminée par des processus se déroulant sur les frontières des grains monocristallins et non pas par des processus dans le volume du semiconducteur; dans un tel cas la relation (9.17) n'est plus applicable.

§ 9.4. Phénomènes acousto-électroniques

Interactions des électrons avec des vibrations acoustiques de faibles fréquences du réseau. Examinons maintenant l'interaction entre les électrons et les vibrations acoustiques de faibles fréquences

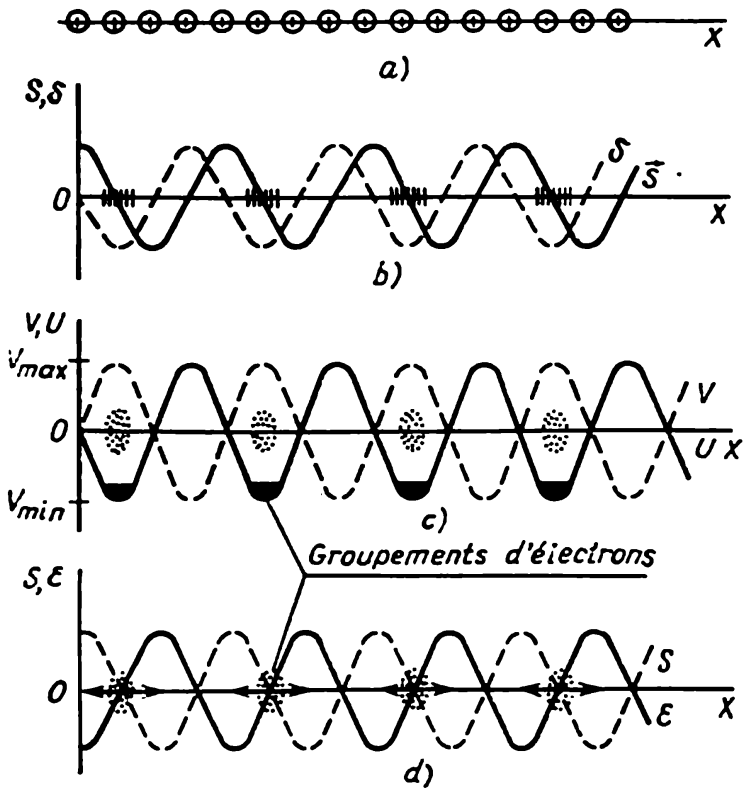

Fig. 9.12. Interaction des ondes élastiques avec les électrons libres d'un cristal: *a* — chaîne d'ions de métal chargés positivement; *b* — déplacement des ions *S* et variation de la distance relative entre les ions δ lors de la propagation de l'onde élastique (l'hachure verticale désigne les régions de compression maximale de la chaîne); *c* — oscillation du potentiel *V* de champ et de l'énergie potentielle *U*; *d* — oscillation de l'intensité de champ électirque $\mathscr{E}$; les flèches indiquent le sens du champ.

excitées dans le cristal. Le mécanisme de cette interaction peut être différent suivant la nature du cristal.

Considérons à titre d'exemple le plus simple un cristal métallique unidimensionnel — une chaîne d'ions chargés positivement plongée dans un gaz électronique (fig. 9.12, *a*). Excitons dans ce cristal une onde progressive longitudinale

$$s_x = A \cos 2\pi (t/T - x/\lambda), \tag{9.20}$$

où s_x est le déplacement de l'ion se trouvant à la distance x de l'origine des coordonnées; λ la longueur d'onde; T la période; A l'amplitude de vibrations. Le déplacement des ions se produit le long de l'axe des x. La figure 9.12, *b* montre en traits pleins le déplacement s des

ions à un instant fixé t (pour simplifier nous posons t égal à zéro). La courbe en traits interrompus indique la variation relative de la distance entre les ions

$$\delta = \frac{ds}{dx} = A\frac{2\pi}{\lambda}\sin 2\pi\left(\frac{t}{T} - \frac{x}{\lambda}\right). \qquad (9.21)$$

La valeur négative maximale de δ est atteinte aux points $x = \lambda/4$, $5(\lambda/4)$, . . . En ces points la chaîne est serrée et la distance entre les ions est inférieure à la distance à l'état d'équilibre. La valeur positive maximale de δ est atteinte aux points $x = 3(\lambda/4)$, $7(\lambda/4)$, . . . En ces endroits, le réseau est étiré au maximum et la distance entre les ions est plus grande qu'à l'état d'équilibre. La variation de la densité de répartition des ions le long de la chaîne, lorsqu'on y excite une onde longitudinale, provoque une augmentation de la densité de la charge positive aux endroits d'épaississement de l'onde et une diminution de cette densité aux endroits de raréfaction. Il en résulte l'apparition d'un champ électrique dont le potentiel positif est maximal aux épaississements de l'onde et minimal aux raréfactions. Le caractère de variation du potentiel V le long de la chaîne à un instant fixé est indiqué sur la figure 9.12, *c* par la courbe en traits interrompus et la variation de l'énergie potentielle de l'électron dans le champ de l'onde ($U = -eV$) est indiquée en traits pleins. La figure 9.12, *d* représente la courbe d'intensité du champ électrique de l'onde : $\mathscr{E} = -\frac{dV}{dx}$ (courbe en traits pleins). Cette intensité est maximale aux endroits où les déplacements s sont minimaux et, inversement, elle est minimale là où s est maximal ; le sens de $\mathscr{E}$ est indiqué par des flèches horizontales.

Ainsi, une onde sonore progressive se propageant dans un métal fait naître une onde électrique animée de la même vitesse que l'onde sonore. La mise sous écran de ce champ se fait par redistribution des électrons libres : aux endroits de minimum d'énergie potentielle la concentration des électrons augmente et aux endroits de maximum elle diminue. Les électrons « s'écoulent » dans les puits de potentiel en formant des amas électroniques. La propagation de l'onde fait déplacer les puits de potentiel le long de la chaîne avec la vitesse du son v_{son} en entraînant les électrons qu'ils renferment. C'est la raison pour laquelle on donne à cet effet le nom d'*effet d'entraînement des électrons par des phonons acoustiques de faibles fréquences.*

Un phénomène analogue s'observe aussi lors de la propagation des ondes acoustiques de faibles fréquences dans des cristaux tels que le germanium, le silicium et autres, mais le mécanisme de son apparition est différent. La compression et l'extension du réseau cristallin font varier les distances entre les atomes, ce qui conduit à une variation de la largeur de la bande interdite E_g. Notamment, pour le germanium et le silicium l'augmentation de la pression

de compression extérieure p fait croître la largeur de la bande interdite à peu près proportionnellement à la pression : $\Delta E_g = \beta p$, où β est un coefficient de proportionnalité appelé *constante de potentiel de déformation*. Pour le germanium et le silicium il est relativement faible et une variation appréciable de E_g ne s'observe que dans le cas des pressions relativement fortes.

Le passage d'une onde élastique longitudinale à travers le cristal a pour effet de provoquer l'apparition d'endroits de compression où la largeur de la bande interdite augmente un peu et d'endroits d'extension où E_g diminue légèrement par rapport à un cristal non

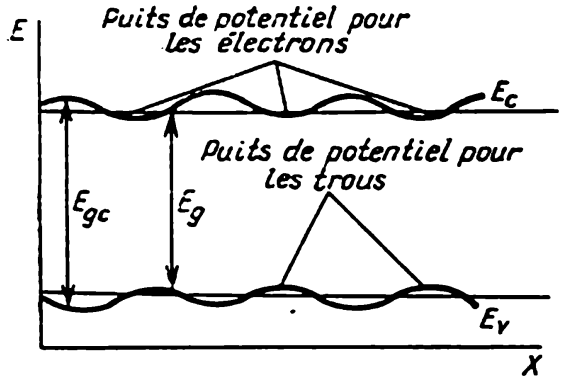

Fig. 9.13. Variation de la largeur de la bande interdite d'un semiconducteur sous l'action d'une onde élastique : E_g — largeur de la bande interdite dans un cristal non déformé ; E_{gc} — largeur de la bande interdite aux endroits de compression du réseau cristallin.

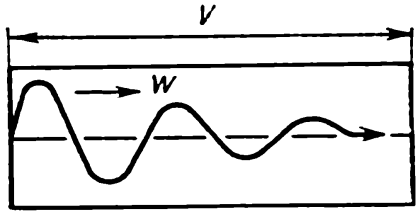

Fig. 9.14. Effet acousto-électrique.

déformé. Il en résulte que l'énergie au niveau inférieur de la bande de conduction et au sommet de la bande de valence varie avec une période égale à la longueur de l'onde sonore comme le montre schématiquement la figure 9.13. Ainsi, il se forme dans l'espace des puits de potentiel, périodiquement répartis, pour les électrons (et les trous), ce qui conduit à une redistribution des électrons et à l'augmentation de leur concentration aux endroits où se situent les puits de potentiel. Lors de la propagation de l'onde élastique le long du cristal les puits de potentiel se déplacent en entraînant une partie des électrons.

Une interaction particulièrement forte des électrons avec les phonons de faibles fréquences se produit dans des cristaux semiconducteurs présentant des propriétés piézo-électriques (dans CdS, CdSe, InSb et beaucoup d'autres). Comme la déformation élastique des cristaux piézo-électriques provoque leur polarisation, la propagation d'une onde sonore longitudinale dans le cristal s'accompagne d'une apparition d'une onde électrique qui se propage à la vitesse du son et entraîne les électrons.

Effet acousto-électrique. L'entraînement des électrons par les ondes sonores donne lieu à un phénomène important qui a reçu le nom d'*effet acousto-électrique*.

Excitons dans un cristal piézo-électrique de type n une onde acoustique (fig. 9.14). En se propageant dans le cristal, cette onde s'affaiblit progressivement, en cédant son énergie aux électrons qu'elle entraîne (nous ne considérons pas l'amortissement de l'onde par suite de son absorption par le réseau cristallin). En plus de l'énergie, l'onde transmet aux électrons son impulsion mécanique. Aussi

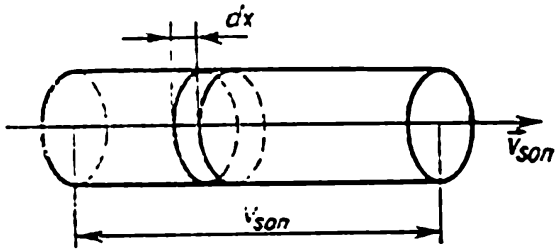

Fig. 9.15. Etablissement des relations pour la tension et le courant acousto-électriques.

les électrons seront-ils soumis, dans le sens de propagation de l'onde, à une certaine force F. Cette force fait déplacer les électrons de la partie gauche du cristal dans la partie droite (fig. 9.14). Il en résulte l'apparition d'un champ électrique qui conduit à la dérive des électrons de droite à gauche. A l'état stationnaire, le circuit étant ouvert, l'intensité de champ $\mathcal{E}$ doit être telle qu'à chaque section du cristal le flux d'électrons entraîné par l'onde sonore soit compensé par le flux de dérive opposé.

Nous considérerons l'onde comme un flux de phonons d'énergie $\hbar\omega$ et d'impulsion $\hbar\omega/v_{son}$ (ω étant la fréquence cyclique des vibrations des atomes dans l'onde). Isolons, par la pensée, dans le cristal, dans le sens de propagation de l'onde, un cylindre dont la base a une surface unité et la longueur est égale à v_{son} (fig. 9.15). Tous les phonons contenus dans ce cylindre passeront pendant 1 s par la base S et transporteront une énergie

$$W = n_{phon} v_{son} \hbar\omega \tag{9.22}$$

et une impulsion

$$p = W/v_{son}, \tag{9.23}$$

où n_{phon} est la concentration des phonons. Dans une couche de cristal d'épaisseur dx est absorbée (fig. 9.15) chaque seconde une énergie $W\gamma dx$, où γ est le coefficient d'absorption de l'énergie de l'onde par les électrons. Les électrons situés dans cette couche reçoivent suivant la relation (9.23) une impulsion égale à $\gamma W dx/v_{son}$. Suivant le second principe de la dynamique, la variation d'impulsion

en unité de temps est égale à la force moyenne s'exerçant sur les électrons de la part de l'onde sonore:

$$\gamma \frac{W\,dx}{v_{son}} = Fn_0\,dx.$$

D'après ce qui précède cette force doit être équilibrée par la force $-e\mathscr{E}n_0dx$ qui agit sur les électrons de la couche dx de la part du champ électrique produit dans le cristal. Ainsi,

$$\gamma \frac{W}{v_{son}}\,dx - e\mathscr{E}n_0\,dx = 0. \tag{9.24}$$

On en trouve pour $\mathscr{E}$:

$$\mathscr{E} = \frac{\gamma W}{en_0 v_{son}} = \frac{u\gamma W}{\sigma v_{son}}, \tag{9.25}$$

où u est la mobilité des électrons, σ la conductivité électrique du cristal. Entre les extrémités du cristal de longueur L il apparaît une différence de potentiel

$$V = -\int_0^L \mathscr{E}\,dx = -\frac{u\gamma W}{\sigma v_{son}} L. \tag{9.26}$$

Si les extrémités du cristal sont reliées l'une à l'autre, le circuit ainsi formé sera parcouru par un courant

$$I = V/R = -u\gamma WS/v_{son}, \tag{9.27}$$

où $R = \rho \frac{L}{S}$ est la résistance du cristal dont l'aire de la section droite est égale à S. La densité de courant est

$$i = -\frac{u\gamma W}{v_{son}}. \tag{9.28}$$

Ainsi, *l'effet d'entraînement des électrons par l'onde sonore conduit à l'apparition d'une différence de potentiel entre les extrémités d'un cristal piézo-électrique ouvert et au passage d'un courant I dans le circuit fermé du cristal.* Ce phénomène a reçu le nom d'effet acousto-électrique; les grandeurs V et I sont appelées respectivement *tension acousto-électrique* et *courant acousto-électrique.*

Amplification d'ondes ultra-sonores. Nous avons étudié l'interaction des électrons avec une onde élastique dans le cas où le champ électrique extérieur dans le cristal est absent. Imaginons maintenant qu'une onde acoustique se propage dans un cristal piézo-électrique dans lequel on a créé un champ électrique extérieur qui provoque la dérive des électrons à la vitesse v_{der}, le sens de propagation de l'onde et le sens de mouvement de dérive étant confondus. Dans ce cas lui aussi, les électrons interagissent avec l'onde élastique, mais le

résultat de cette interaction dépend essentiellement du rapport qui existe entre la vitesse de dérive v_{der} et la vitesse de l'onde v_{son}. La fig. 9.16 montre la variation du coefficient d'absorption électronique du son en fonction du rapport v_{der}/v_{son}. Pour $v_{der} = 0$ (le champ extérieur est nul) le coefficient d'absorption a une certaine valeur γ_0. Lorsque la vitesse v_{der} augmente, le coefficient d'absorption γ s'accroît d'abord et, pour une certaine valeur du rapport v_{der}/v_{son}, il prend sa valeur maximale

$$\gamma_{max} = C\omega, \qquad (9.29)$$

où C est un coefficient de proportionnalité dépendant des paramètres mécaniques et électriques du matériau ; ω la fréquence. Si la vitesse v_{der} croît encore, le coefficient γ commence à diminuer rapidement, s'annule pour $v_{der}/v_{son} = 1$ et devient négatif quand $v_{der}/v_{son} > 1$.

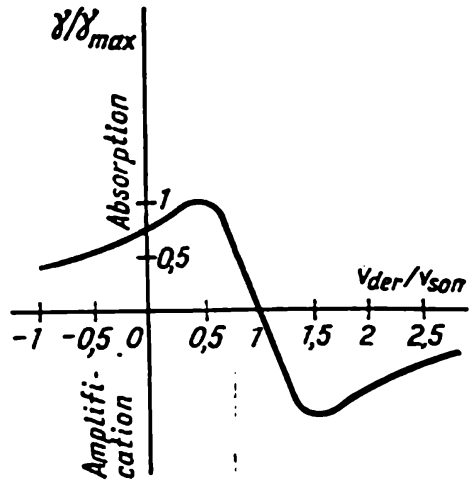

Fig. 9.16. Variation du coefficient d'absorption électronique du son en fonction du rapport de la vitesse de dérive v_{der} des électrons à la vitesse du son v_{son}.

Physiquement on peut en donner l'interprétation suivante. Si la vitesse de dérive v_{der} des électrons est égale à la vitesse de l'onde sonore, les électrons n'interagissent pas avec l'onde parce qu'étant réunis en groupements dans les puits de potentiel (fig. 9. 17, *a*), ils se déplacent avec l'onde. Le coefficient d'absorption électronique du son est dans ce cas nul. Pour $v_{der} < v_{son}$, les électrons retardent par rapport à l'onde sonore et donc par rapport au champ électrique engendré par cette onde dans le cristal. Ce champ commence à tirer (accélérer) les groupements d'électrons en transmettant aux électrons l'énergie de l'onde sonore. L'onde s'amortit (fig. 9. 17, *b*), ce qui correspond à une valeur positive du coefficient d'absorption électronique du son. Quand $v_{der} > v_{son}$ les électrons avancent sur l'onde sonore de sorte que le champ électrique produit par l'onde tend à freiner les groupements d'électrons, en leur prélevant une partie d'énergie qu'il transmet à l'onde sonore. Le coefficient d'absorption électronique γ aura dans ce cas une valeur négative, ce qui correspond à l'effet d'amplification de l'onde sonore au détriment de l'énergie des électrons en dérive ou plus exactement au détriment de l'énergie de la source électrique qui entretient la dérive des électrons ; l'amplitude de l'onde sonore s'en trouve augmentée (fig. 9.17, *c*). Pour des ondes de fréquences acoustiques cet effet est relativement faible. Il augmente avec l'augmentation de la fréquence ω [formule (9.29)] et atteint des valeurs notables pour des ondes ultrasonores ($\nu \approx 1$

à 10 MHz). C'est ainsi que $\gamma_{max} \approx 18$ cm^{-1} pour le CdS à la fréquence de l'onde $\nu = 30$ MHz ; lorsque cette onde se propage dans une plaque de 0,5 cm de longueur son intensité devient $e^{0,5 \cdot 18} \approx 10^4$ fois plus grande. Comme le montre la figure 9.16 cette amplification sera observée pour $v_{der} = 1,5 v_{son}$. Pour obtenir une telle vitesse de

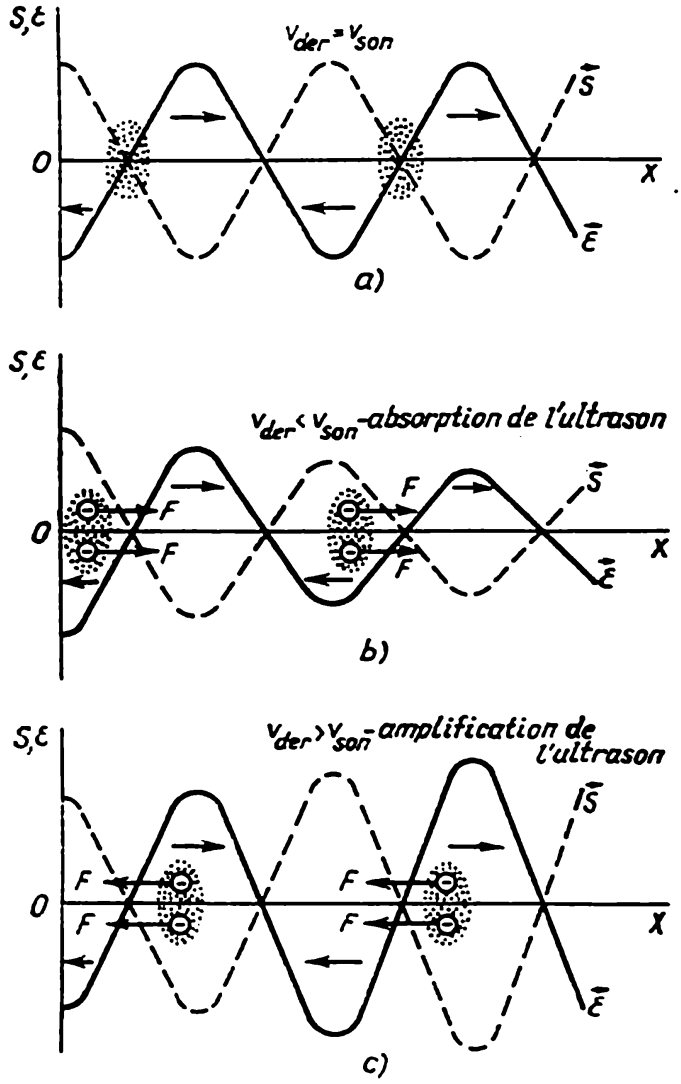

Fig. 9.17. Interaction des électrons avec une onde sonore pour différents rapports de la vitesse de dérive des électrons à la vitesse de propagation de l'onde (la flèche représente la force qui agit sur les électrons de la part de l'onde).

dérive des électrons il faut appliquer au cristal de CdS un champ extérieur d'intensité

$$\mathscr{E} = \frac{v_{der}}{u} = \frac{1,5 v_{son}}{u} = \frac{1,5 \cdot 4,3 \cdot 10^5 \text{ cm/s}}{200 \text{ cm}^2/(\text{V} \cdot \text{s})} \approx 3,2 \text{ kV/cm},$$

où $v_{son} = 4,3 \cdot 10^5$ cm/s est la vitesse des ondes longitudinales dans le CdS et $u = 200$ cm^2/(V·s) la mobilité des électrons dans le CdS.

La figure 9.18 représente le schéma de principe d'une installation expérimentale pour l'étude de l'absorption électronique et de l'amplification de l'ultrason. L'échantillon *Ech* comportant des électrodes métalliques E_1 et E_2 est serré entre deux convertisseurs électromécaniques C_1 et C_2 constitués par des plaquettes de quartz taillées dans le sens piézo-actif et recouvertes en bas et en haut par des couches métalliques. Le générateur *G* fournit des oscillations électriques de fréquence égale à la fréquence des oscillations propres de la plaquette C_1. Le convertisseur C_1 les transforme en oscillations ultrasonores (effet piézo-électrique inverse) qui se propagent sous forme d'ondes ultrasonores dans l'échantillon *Ech*. Ces ondes sont reçues par le convertisseur C_2 qui les convertit en oscillations électriques enregistrées par l'oscillographe électronique *OE*. Entre les électrodes E_1 et E_2 est appliquée une tension produisant dans l'échantillon un champ d'entraînement $\mathscr{E}_0$ qui provoque la dérive des électrons dans le sens de propagation de l'onde. Pour éviter un échauffement anormal de l'échantillon, le champ $\mathscr{E}_0$ est produit par des impulsions de tension de forme rectangulaire et de courte durée. En faisant varier l'intensité du champ $\mathscr{E}_0$ et donc la vitesse de dérive des électrons $v_{der} = u\mathscr{E}_0$, on peut parcourir toute la courbe de variation du coefficient d'absorption électronique γ en fonction du rapport v_{der}/v_{son}.

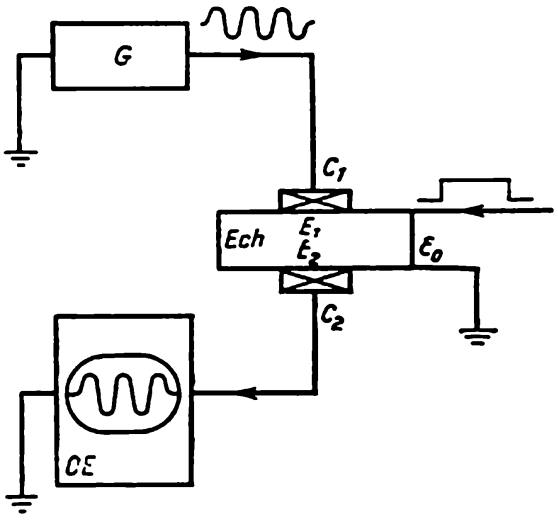

Fig. 9.18. Schéma de principe d'une installation pour l'étude de l'absorption et l'amplification électroniques des ondes sonores.

Les expériences montrent que pour $v_{der} < v_{son}$ il se produit réellement une absorption électronique intense du son et l'amplitude de signal enregistrée à la sortie par l'oscillographe peut être beaucoup plus faible que l'amplitude des oscillations appliquées au convertisseur C_1. Au contraire, pour des vitesses de dérive $v_{der} > v_{son}$ on observe l'amplification des ondes ultrasonores de sorte que le signal de sortie peut être beaucoup plus intense que le signal injecté dans le convertisseur C_1.

Comme il a été indiqué plus haut, l'effet d'entraînement des électrons par les ondes sonores a lieu aussi dans des semiconducteurs non piézo-électriques. Aussi, dans de tels semiconducteurs peut-on en principe obtenir l'amplification de l'ultrason. Toutefois, étant donné que l'interaction des électrons avec l'onde sonore élastique est beaucoup plus faible dans les semiconducteurs non piézo-électri-

ques que dans les semiconducteurs piézo-électriques, l'effet d'amplification dans les premiers est beaucoup moins intense et ne s'observe qu'à des fréquences très élevées des ondes sonores (hyperson).

Ondes acoustiques de surface. C'est pour le cas des *ondes acoustiques de surface* (OAS) que les phénomènes qui viennent d'être considérés présentent un intérêt particulier et une importance pratique. De telles ondes se concentrent directement près de la surface d'un cristal piézo-électrique dans les limites de quelques longueurs d'onde (pour des ondes de fréquence $\nu \approx 300$ MHz la profondeur de pénétration est dans le cas du quartz d'environ 30 μm). Pour la production de ces ondes on utilise des convertisseurs électromécaniques à tiges

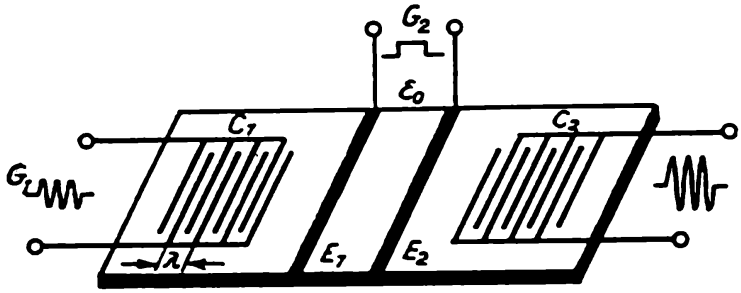

Fig. 9.19. Schéma de principe d'une installation pour l'amplification des ondes acoustiques de surface.

opposées (en forme de peigne) C_1 et C_2 représentés schématiquement sur la figure 9.19. Ces convertisseurs sont des bandes métalliques appliquées sur la surface du cristal piézo-électrique et représentant une sorte de deux peignes qui s'emboîtent l'un dans l'autre. Ces peignes sont attaqués, depuis un générateur d'oscillations électriques G_1, par une tension alternative qui excite dans le cristal piézo-électrique une onde acoustique de surface. La fréquence de la tension alternative doit être telle que la longueur de l'onde excitée soit égale à la période du peigne. La transformation des vibrations mécaniques de l'onde sonore en oscillations électriques se fait par un même convertisseur électromécanique C_2.

Pour assurer l'amplification des ondes acoustiques de surface grâce au mouvement de dérive des électrons, on porte sur la surface du cristal piézo-électrique deux électrodes E_1 et E_2 réalisées sous forme de deux bandes métalliques parallèles auxquelles on applique une tension impulsionnelle fournie par le générateur G_2. Cette tension produit un champ d'entraînement qui provoque la dérive des électrons dans le sens de propagation de l'onde acoustique de surface. Pour $v_{dcr} > v_{son}$ on observe l'effet d'amplification de ces ondes.

Les phénomènes acousto-électroniques utilisant des ondes acoustiques de surface ont permis de réaliser certains appareils et dispositifs. Une faible vitesse de propagation des OAS par rapport à la vitesse de la lumière ($v_{son} \approx 10^{-5}$ s) a permis de réaliser des *lignes à retard*

(LR) de faible encombrement à ondes acoustiques de surface qui trouvent des applications dans les systèmes de radar, la télévision, les calculateurs numériques, etc. La fréquence de service d'une ligne à retard se détermine par le pas des électrodes du convertisseur électromécanique et son temps de retard par la distance entre les convertisseurs d'entrée et de sortie, la bande passante par l'étendue (le nombre de tiges) du convertisseur. Dans la gamme d'ultra-hautes fréquences la longueur de l'onde acoustique de surface est de $\approx 20\,\mu$m. Une ligne à retard de 1 cm de longueur utilisant les ondes acoustiques de surface permet d'obtenir le même retard qu'un câble coaxial long de 1 km. Chose importante, les pertes dans une ligne à retard à ondes acoustiques de surface peuvent être compensées au détriment de l'énergie des électrons en dérive.

Les lignes à retard dispersives, dont le temps de retard varie linéairement avec la fréquence, ont permis de réaliser des filtres de compression et d'expansion de signaux ayant un coefficient de compression allant jusqu'à 10^4.

Une double conversion — signal électrique — onde sonore et onde sonore — signal électrique — sert de base pour la réalisation d'amplificateurs et d'oscillateurs y compris à fréquence accordable.

§ 9.5. Effet Gunn

Apparition d'une conductance différentielle négative dans les semiconducteurs homogènes sous l'effet d'un champ intense. Dans des champs électriques de forte intensité, la mobilité des porteurs de charges commence à dépendre de l'intensité du champ appliqué: $u = u(\mathscr{E})$. Il s'ensuit que tout en restant positive, la conductivité statique du semiconducteur $\sigma_0 = enu$ figurant dans la loi d'Ohm ($i = enu\,\mathscr{E} = \sigma_0\mathscr{E}$) peut varier notablement avec l'intensité du champ. Suivant le caractère de cette variation, la conductance différentielle du semiconducteur

$$\sigma_{dif} = \frac{di}{d\mathscr{E}} = enu + en\mathscr{E}\frac{du}{d\mathscr{E}} = \sigma_0\left(1 + \frac{\mathscr{E}}{u}\frac{du}{d\mathscr{E}}\right) \qquad (9.30)$$

peut prendre une valeur aussi bien positive que négative.

Le premier cas se réalise lorsque l'augmentation de l'intensité de champ $\mathscr{E}$ provoque une augmentation de la mobilité u des porteurs de telle sorte que $\frac{\mathscr{E}}{u}\frac{du}{d\mathscr{E}} > 0$, ou sa diminution mais si faible que bien que $\frac{\mathscr{E}}{u}\frac{du}{d\mathscr{E}} < 0$, la valeur absolue $\left|\frac{\mathscr{E}}{u}\frac{du}{d\mathscr{E}}\right| < 1$, de sorte que l'expression entre parenthèses dans (9.30) conserve son signe positif.

Le second cas se réalise lorsque la croissance de $\mathscr{E}$ entraîne une diminution de la mobilité des porteurs de charges et cette diminution est si brusque que se trouve vérifiée non seulement la condition

$\frac{\mathscr{E}}{u}\frac{du}{d\mathscr{E}} < 0$, mais aussi la condition $\left|\frac{\mathscr{E}}{u}\frac{du}{d\mathscr{E}}\right| > 1$. Alors l'expression entre parenthèses dans (9.30) devient négative, ce qui conduit à une valeur négative de σ_{diff}.

Réalisation de la conductance différentielle négative sous l'effet d'un champ intense dans l'arséniure de gallium. Le cas d'apparition d'une conductance différentielle négative sous l'effet d'un champ électrique de forte intensité que nous venons de considérer se réalise facilement dans les cristaux d'arséniure de gallium. La figure 9.20

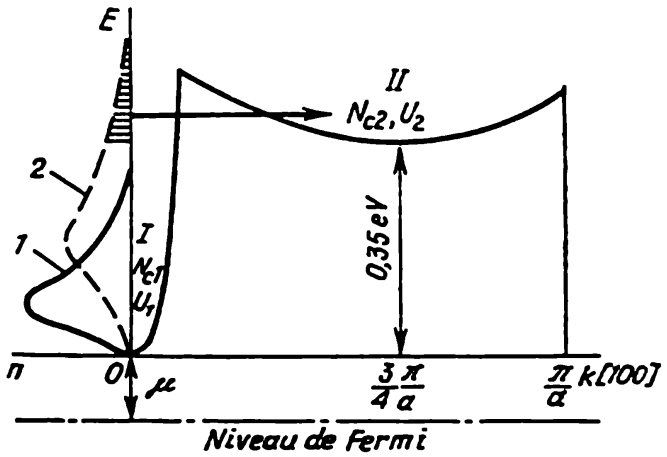

Fig. 9.20. Diagramme énergétique de la bande de conduction de l'arséniure de gallium.

(partie droite) montre la structure énergétique de la bande de conduction de l'arséniure de gallium. Cette bande présente deux minimums: le premier (I) se situe au milieu de la zone de Brillouin et le second (II) à une distance de $^3/_4\ \pi/a$ du centre de la bande. Suivant l'axe des énergies le second minimum se situe de $\sim$ 0,35 eV plus haut que le premier. La densité d'états N_{c1} dans le minimum inférieur est d'environ 1500 fois plus faible que dans le minimum supérieur N_{c2} (c'est-à-dire que $N_{c2}/N_{c1} \approx 1500$), la mobilité u des électrons dans le minimum inférieur est d'environ 40 fois plus grande que dans le minimum supérieur ($u_1/u_2 \approx 40$).

En l'absence de champ électrique extérieur les électrons ayant passé des niveaux donneurs dans la bande de conduction, sont en équilibre thermodynamique avec le réseau du semiconducteur et ont la même temérature T_0 que le réseau. Ils peuvent occuper des niveaux énergétiques aussi bien dans le minimum inférieur que dans le minimum supérieur de cette bande. Leurs concentrations dans ces minimums ont pour valeurs respectives

$$n_1 = N_{c1}e^{\mu/(kT_0)} \quad \text{et} \quad n_2 = N_{c2}e^{(\mu-0,35)/(kT_0)},$$

où μ est la distance entre le niveau inférieur de la bande de conduction et le niveau de Fermi (v. fig. 9.20). Le rapport

$$\frac{n_2}{n_1} = \frac{N_{c2}}{N_{c1}} e^{-0,35/(kT_0)} \approx 1500 e^{-0,35/(kT_0)}.$$

Pour $T_0 = 300$ K on a $kT_0 = 0,026$ eV et $n_2/n_1 = 2,5 \cdot 10^{-3}$.

Ainsi, bien que statistiquement il est plus avantageux pour les électrons de se trouver dans le minimum supérieur présentant une plus grande densité d'états, néanmoins pour $T_0 = 300$ K ils ne peuvent pas l'atteindre parce que leur énergie n'est pas suffisante pour cela. Aussi, en pratique ils se situent presque tous ($\sim$ 99,8 %) dans le minimum inférieur.

La partie gauche de la figure 9.20 montre la courbe de distribution de Boltzmann des électrons de la bande de conduction suivant les énergies à $T = 300$ K (courbe *1*). Elle ne s'étend pratiquement pas dans le domaine des énergies correspondant au minimum supérieur.

Examinons maintenant l'influence qu'un champ électrique fort peut avoir sur le caractère de distribution des électrons suivant les énergies. En provoquant un mouvement de dérive des électrons, le champ électrique leur transmet de l'énergie. Par suite de la diffusion des électrons, cette énergie se transforme en énergie de leur agitation thermique désordonnée — le gaz électronique « s'échauffe ». Dans un champ intense sa température T_e peut dépasser notablement la température T_0 du réseau cristallin. Il en résulte une augmentation de l'énergie des électrons et la courbe de leur distribution suivant les énergies se déforme en s'étendant dans le domaine des plus hautes énergies (fig. 9.20, courbe *2*). Cela conduit à ce que le nombre des électrons capables de passer du minimum inférieur au minimum supérieur devient toujours plus grand. Sur la figure 9.20 la partie de la courbe de distribution correspondant à l'occupation du minimum supérieur est hachurée et les transitions elles-mêmes des électrons du minimum *I* au minimum *II* sont indiquées par une flèche horizontale.

Un calcul élémentaire montre que pour l'arséniure de gallium

$$T_e \sim \mathscr{E}^{\gamma}, \tag{9.31}$$

et pour des champs à partir de $\mathscr{E} \approx 3 \cdot 10^5$ V/m le coefficient $\gamma > 1$. C'est pourquoi dans des champs d'intensité $\mathscr{E} > 3 \cdot 10^5$ V/m la température du gaz électronique commence à croître brutalement avec $\mathscr{E}$ et atteint déjà une valeur de $T_e \approx 600$ K pour $\mathscr{E} \approx 3,5 \cdot 10^5$ V/m. A une telle température du gaz électronique le rapport $n_2/n_1 \approx 1,75$. Cela signifie que pour $\mathscr{E} \approx 3,5 \cdot 10^5$ V/m la plus grande partie des électrons de la bande de conduction sont concentrés dans le minimum supérieur et non inférieur. La mobilité des électrons étant beaucoup plus faible (de 40 fois) dans le minimum supérieur que dans le mini-

mum inférieur, le passage d'un grand nombre d'électrons du minimum inférieur au minimum supérieur doit s'accompagner d'une diminution brusque de leur mobilité effective et donc d'une diminution de la densité de courant traversant le semiconducteur qui se décrit dans ce cas par la formule suivante:

$$i = e\,(n_1 u_1 + n_2 u_2)\mathscr{E}. \tag{9.32}$$

Sur la figure 9.21 la droite OD montre la dépendance $i_1(\mathscr{E})$ tracée dans l'hypothèse où tous les électrons de la bande de conduction se trouvent au minimum inférieur ($n_1 = n$, $n_2 = 0$) et possèdent la mobilité u_1. Cette droite est inclinée de l'angle α_1 sur l'axe des abscisses. Sur la même figure est tracée la droite OC représentative

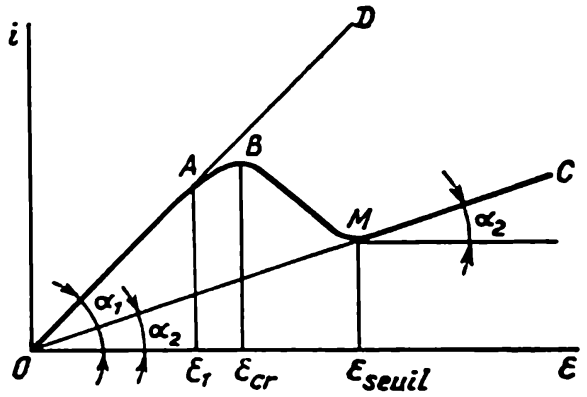

Fig. 9.21. Caractéristique courant-tension en N d'un semiconducteur à résistance différentielle négative produite sous l'action d'un champ électrique intense.

de la dépendance $i_2(\mathscr{E})$ dans l'hypothèse que tous les électrons se situent au minimum supérieur ($n_1 = 0$, $n_2 = n$) et possèdent la mobilité u_2. Cette droite est inclinée sur l'axe des abscisses d'un angle $\alpha_2 < \alpha_1$.

Voyons maintenant comment varie la densité de courant dans le semiconducteur au fur et à mesure que l'intensité de champ $\mathscr{E}$ augmente. Tant que l'intensité $\mathscr{E}$ est insuffisante pour assurer un échauffement notable des électrons, tous les électrons restent au minimum inférieur et la dépendance $i(\mathscr{E})$ est décrite par la droite OD. Toutefois à mesure que $\mathscr{E}$ augmente, un nombre toujours plus grand d'électrons acquièrent l'énergie suffisante pour passer du minimum inférieur au minimum supérieur. Comme cette transition s'accompagne d'une chute de la mobilité des électrons, elle conduit à une diminution de la densité de courant. C'est pourquoi à partir d'une certaine intensité $\mathscr{E}_1$ la croissance du courant i avec la croissance de $\mathscr{E}$ commence à se ralentir pour s'arrêter complètement pour $\mathscr{E} = \mathscr{E}_{cr}$. Si l'intensité $\mathscr{E}$ augmente encore, le passage des électrons au minimum supérieur se produit d'une façon si intense que le courant i

non seulement ne reste pas constant mais décroît avec l'augmentation de $\mathscr{E}$ (tronçon BM). Ceci étant, la conductance différentielle du semiconducteur est dans ce tronçon une grandeur négative : $\sigma_{dif} = \frac{di}{d\mathscr{E}} < 0$. La décroissance du courant i avec la croissance de $\mathscr{E}$ continue jusqu'à l'intensité $\mathscr{E}_{seuil}$ pour laquelle la majorité absolue des électrons passent au minimum supérieur. Après cela la dépendance i ($\mathscr{E}$) reprend son caractère linéaire et l'angle d'inclinaison de la droite i ($\mathscr{E}$) sera égal à α_2.

La caractéristique courant-tension d'un tel type comportant une partie à conductance différentielle négative est appelée *caractéristique en* N. Elle présente un grand intérêt pour la radio-électronique vu que les systèmes présentant une telle caractéristique peuvent être utilisés pour l'amplification et la production d'oscillations électromagnétiques ainsi qu'à d'autres fins.

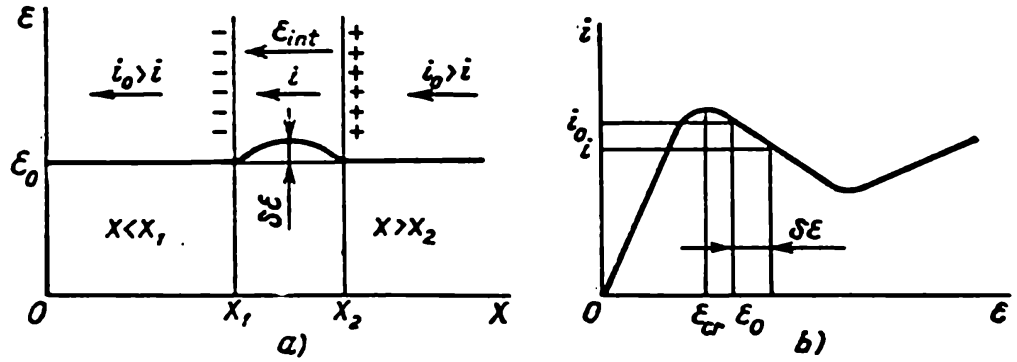

Fig. 9.22. Formation d'un domaine électrostatique dans un semiconducteur à l'endroit de fluctuations du champ électrique dont l'intensité moyenne $\mathscr{E}_0 > \mathscr{E}_{cr}$.

Apparition de domaines électrostatiques. Considérons une propriété commune propre à tous les semiconducteurs qui possèdent dans des champs intenses une caractéristique courant-tension en N.

Supposons qu'un cristal homogène et suffisamment long est parcouru par un courant fourni par une source d'alimentation produisant dans le semiconducteur un champ $\mathscr{E}_0$ tel que $\mathscr{E}_{cr} < \mathscr{E}_0 < \mathscr{E}_{seuil}$ (fig. 9.21). Supposons aussi que ce champ est uniforme de sorte que dans tous les points du cristal l'intensité du champ est la même, comme il est indiqué sur la figure 9.22, *a*. Supposons ensuite que dans un petit tronçon compris entre x_1 et x_2 du cristal le champ a subi une augmentation d'une petite quantité $\delta\mathscr{E}$. Comme le montre la caractéristique courant-tension (fig. 9.22, *b*), la densité de courant dans la région $x_1 < x < x_2$ sera plus faible que dans les régions $x < x_1$ et $x > x_2$. Il en résulte que les électrons commencent à s'accumuler au voisinage de x_1 en y constituant une charge négative et à se détacher des électrons se trouvant pour $x > x_2$ en laissant une

charge positive non compensée. Ainsi il se forme entre les points x_1 et x_2 une couche de dipôle appauvrie en porteurs libres. Cette couche est appelée *domaine électrostatique*. A l'intérieur du domaine il naît un champ interne $\mathscr{E}_{int}$ qui peut dépasser notablement le champ moyen $\mathscr{E}_0$ régnant dans le cristal. Dans un échantillon, alimenté depuis un générateur de tension, la formation de domaine s'accompagne d'une baisse du champ moyen dans le reste du cristal (fig. 9.23)

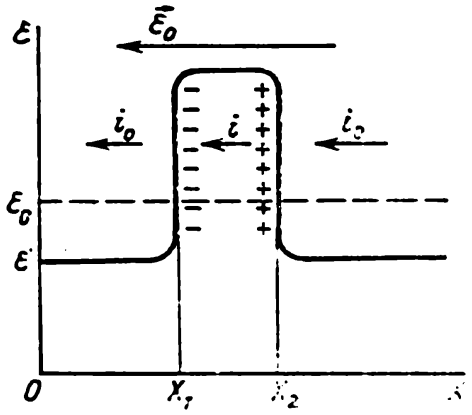

Fig. 9.23. Formation d'un domaine dans un semiconducteur alimenté depuis un générateur de tension s'accompagnant d'une diminution du champ en dehors du domaine.

Déplacement de domaines et apparition d'oscillations de courant. Le processus de formation de domaines décrit plus haut se produit dans une collectivité d'électrons qui se déplacent de l'électrode portée à un potentiel négatif (cathode) vers l'électrode portée à un potentiel positif (anode). C'est pourquoi le domaine se déplace comme un tout dans le sens de la cathode vers l'anode avec une vitesse v_{der} égale pour l'arséniure de gallium à environ 10^5 m/s.

Le champ atteint son intensité critique tout d'abord dans des endroits du cristal où la conductivité a sa plus grande valeur par suite

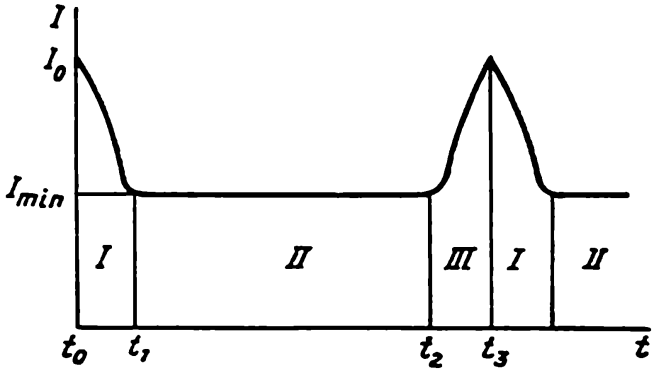

Fig. 9.24. Variation du courant sous forme de « picots » dans un semiconducteur lors de la formation et de la désintégration d'un domaine électrostatique.

des particularités de croissance. En pratique, de tels endroits se situent toujours près des électrodes parce que ces régions du semiconducteur sont les plus inhomogènes par suite de la soudure des contacts.

En prenant naissance près de la cathode, le domaine se déplace vers l'anode où il se désintègre. La désintégration d'un domaine

à l'anode s'accompagne d'une formation d'un autre domaine à la cathode de sorte que le processus prend un caractère périodique de fréquence

$$\nu = v_{der}/L, \tag{9.33}$$

où L est la longueur du cristal. Comme la concentration d'électrons libres dans le domaine est réduite, l'apparition du domaine dans le cristal s'accompagne d'une élévation de la résistance de l'échantillon et d'une réduction de l'intensité du courant dans le circuit.

La figure 9.24 montre le caractère de variation du courant dans le circuit de l'échantillon en fonction du temps. A l'instant t_0 de création du domaine le circuit est parcouru par le courant I_0. Au fur et à mesure que le domaine se forme (intervalle I) le courant diminue pour atteindre sa valeur minimale I_{min} à l'instant t_1 correspondant à la fin du processus de formation du domaine. Dans l'intervalle II de durée $t_2 - t_1 = L/v_{der}$ le domaine se déplace de la cathode vers l'anode et l'intensité de courant dans le circuit reste inchangée et égale à I_{min}. A l'instant t_2 le domaine atteint l'anode et commence à se désintégrer (intervalle III). La désintégration s'achève vers l'instant t_3 et s'accompagne d'une croissance du courant de I_{min} jusqu'à la valeur initiale I_0.

L'effet d'apparition des oscillations de courant dans le circuit d'un semiconducteur lié au passage d'un domaine a été observé expérimentalement pour la première fois dans l'arséniure de gallium de type n par Gunn; il a reçu le nom d'*effet Gunn*. Les appareils utilisant cet effet s'appellent *oscillateurs à effet Gunn* ou *diodes Gunn*.

L'effet Gunn a été découvert en 1963. En 1966 a été construit un premier générateur d'ondes hyperfréquences industriel avec une fréquence de travail de l'ordre de 3 GHz et une puissance de sortie de 100 W en régime d'impulsions. Au point de vue de la construction, les diodes à effet Gunn sont réalisées sous la forme d'une plaquette de cristal homogène d'arséniure de gallium de type n munie de deux contacts non redresseurs. Comme il ressort de la formule (9.33) la fréquence d'oscillations produites par une diode Gunn se détermine par l'épaisseur de la plaquette. La technologie moderne de fabrication de telles plaquettes permet de réaliser des diodes d'épaisseur égale à quelques micromètres, ce qui correspond à la fréquence générée de plusieurs dizaines de gigahertz. La puissance des diodes modernes utilisant l'effet Gunn est de quelques centaines de milliwatts en continu et de quelques centaines de watts en impulsions. La présence d'hétérogénéités dans le cristal a pour résultat que les domaines peuvent se former non seulement près de la cathode mais aussi dans d'autres régions de la plaquette, ce qui explique pourquoi les diodes Gunn et surtout celles qui utilisent des plaquettes relativement épaisses (quelques centaines de micromètres) se caractérisent par un niveau suffisamment élevé de bruits d'amplitude et de fréquence.

CHAPITRE 10

CONTACT MÉTAL-SEMICONDUCTEUR

Dans l'électronique des solides un rôle important est joué par des phénomènes qui se produisent dans les contacts métal-métal, métal-semiconducteur, métal-diélectrique, semiconducteur-semiconducteur, semiconducteur-diélectrique.
Les contacts métal-métal sont utilisés pour la formation de connexions entre les éléments constitutifs des appareils et pour leur branchement sur les circuits extérieurs. Les contacts métal-semiconducteur peuvent être *ohmiques* ou *redresseurs* suivant leur caractère. Les contacts ohmiques sont largement utilisés dans les diodes, les transistors et les éléments semiconducteurs passifs des circuits intégrés. Les propriétés de redressement des contacts métal-semiconducteur sont utilisées pour la réalisation de diodes et de transistors.
Les contacts métal-diélectrique et semiconducteur-diélectrique sont utilisés pour la construction d'appareils à couche mince dont le fonctionnement est fondé sur les courants limités par la charge d'espace (CLCE) et l'effet de champ.
Les contacts métal-semiconducteur et semiconducteur-semiconducteur servent aussi de base pour la construction de générateurs thermo-électriques, d'installations frigorifiques et de nombreux dispositifs semiconducteurs (diodes, transistors, etc.).
Le chapitre actuel se propose d'étudier les propriétés physiques des contacts métamétal, métal-semiconducteur, métal-diélectrique. Pour ce qui est des propriétés du contact semiconducteur-semiconducteur utilisées pour la réalisation d'une large classe d'appareils semiconducteurs elles seront examinées au cours des chapitres qui suivent.

§ 10.1. Travail de sortie

Au chapitre 4 il a été montré que lorsque les atomes se rapprochent et forment des cristaux, les barrières de potentiel pour les électrons, qui séparent les atomes voisins, s'abaissent et se rétrécissent (v. fig. 4.2, *b*). Quant à la barrière de potentiel près de la surface d'un cristal (près de sa frontière extérieure), elle conserve pratiquement la même hauteur que dans des atomes isolés (fig. 10.1, *a*). Tout se passe donc comme si les électrons dans le cristal se trouvaient dans un puits de potentiel dont la sortie exige d'effectuer un travail pour surmonter la force qui s'exerce sur les électrons de la part du cristal. Dans le cas des métaux, l'expression donnant cette force se détermine à partir des considérations suivantes.

Un électron qui s'est échappé du métal et se trouve près de sa surface à une distance x (fig. 10.1, *b*) induit dans le métal une charge $+e$. Cette charge induite agit sur l'électron sorti comme si elle était concentrée sous la surface du métal à la profondeur x en un point qui est symétrique de celui où se trouve l'électron. Cela expli-

que pourquoi la charge induite $+e$ est appelée *image électrique* de la charge $-e$. Elle attire l'électron avec une force

$$F = -e^2/(16\pi\varepsilon_0 x^2), \tag{10.1}$$

que l'on appelle *force d'image électrique.*

A une distance x_0 de la surface du cristal, dont la valeur est de l'ordre de la distance interatomique ou encore plus faible, il est assez

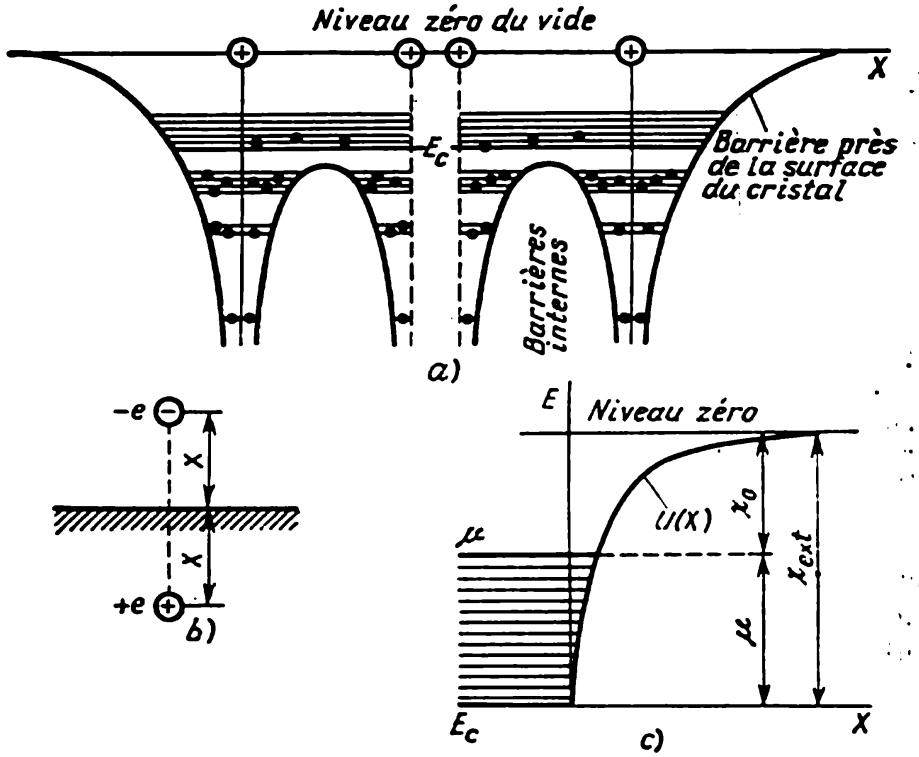

Fig. 10.1. Sortie des électrons d'un métal:

a — barrière de potentiel pour les électrons près de la surface du cristal; *b* — image électrique d'une charge près de la surface d'un métal; *c* — variation de l'énergie potentielle d'un électron $U(x)$ à mesure qu'il s'éloigne de la surface du métal.

difficile de déterminer les forces retenant un électron dans le cristal de sorte que pour des distances $x < x_0$ l'expression (10.1) devient inapplicable. Mais, heureusement, pour la plupart des problèmes ayant une importance pratique, il suffit de connaître seulement la hauteur totale de la barrière comptée à partir du niveau inférieur de la bande de conduction E_c qu'on appelle *travail de sortie* (ou *d'extraction*) *extérieur* χ_{ext} ou encore *affinité électronique* (fig. 10.1, *c*); la hauteur de la barrière comptée à partir du niveau de Fermi μ qui est appelée *travail de sortie thermodynamique* χ_0 (fig. 10.1, *c*) et enfin le potentiel de la force d'image reflétée pour $x > x_0$ qui peut être déterminé par intégration de l'expression (10.1).

Pour un semiconducteur ou un diélectrique (fig. 10.2), le travail de sortie thermodynamique χ_0 ne correspond pas au travail de sortie

d'un électron réel quelconque si le niveau de Fermi se situe dans la bande interdite et ne coïncide avec aucun niveau d'impureté.

Le travail de sortie est fortement influencé par des couches d'adsorption monomoléculaires. La figure 10.3, *a* montre conventionnellement une couche monoatomique de césium déposée sur la surface de tungstène. Le césium est un métal alcalin. Son électron superficiel de valence est lié au noyau beaucoup plus faiblement que les électrons de valence dans le tungstène. Aussi, lors de l'adsorption, les atomes de césium cèdent-ils leurs électrons de valence au tungstène et se transforment en ions chargés positivement. Entre ces ions et leur image électrique dans le tungstène prend naissance une force d'attraction qui maintient les ions sur la surface de tungstène. On peut en conclure que si l'on porte sur la surface de tungstène une couche monoatomique de césium, il se forme une double couche d'électricité dont la face extérieure est chargée positivement. Le champ de la double couche favorise la sortie de l'électron du tungstène si bien qu'en présence de couche de césium le travail de sortie de

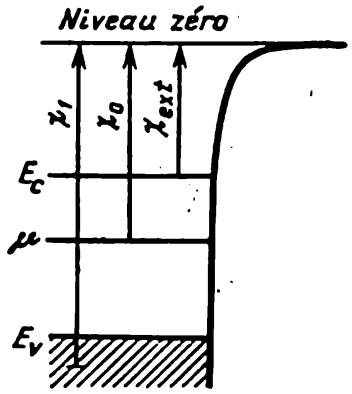

Fig. 10.2. Sortie des électrons d'un semiconducteur:

χ_0 — travail de sortie thermodynamique; χ_{ext} — travail de sortie extérieur; χ_i — travail de sortie d'un électron de l'intérieur de la bande de valence.

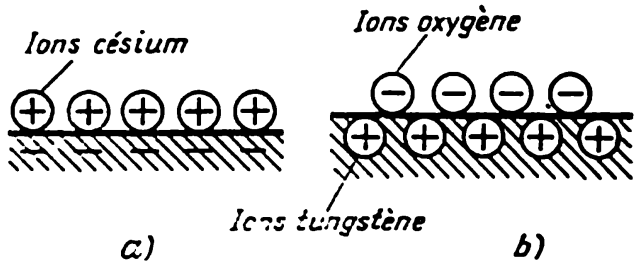

Fig. 10.3. Formation d'une double couche d'électricité lors du dépôt d'une couche de césium sur la surface d'un cristal (*a*) et lors de l'adsorption des atomes d'oxygène sur la surface de tungstène (*b*).

l'électron du tungstène diminue de 4,52 à 1,36 eV. De même que le césium, agissent aussi les couches monoatomiques des autres métaux électropositifs — du baryum, du cérium, du thorium, etc.

La diminution du travail de sortie sous l'effet d'adsorption des éléments électropositifs est largement utilisée en pratique pour la fabrication de cathodes des tubes électroniques, de photocathodes, etc.

Toute différente est l'action de l'oxygène, par exemple, adsorbé par la surface d'un métal. La liaison des électrons de valence dans l'atome d'oxygène est beaucoup plus forte que dans les métaux. Aussi lors de l'adsorption, l'atome d'oxygène ne cède-t-il pas mais, au contraire, reçoit un électron du métal en se transformant en un ion chargé négativement. Il en résulte que la face extérieure de la double couche d'électricité se trouve chargée négativement (fig. 10.3, *b*) et le champ électrique freine donc la sortie des électrons du métal de sorte que le travail de sortie augmente.

§ 10.2. Emission thermoélectronique

Courant d'émission; formule de Richardson-Dushmann. La figure 10.4 montre le diagramme des énergies du tungstène et la courbe de distribution des électrons suivant les énergies à $T = 0$ K et à une température élevée. Comme il est visible sur cette figure, lorsque

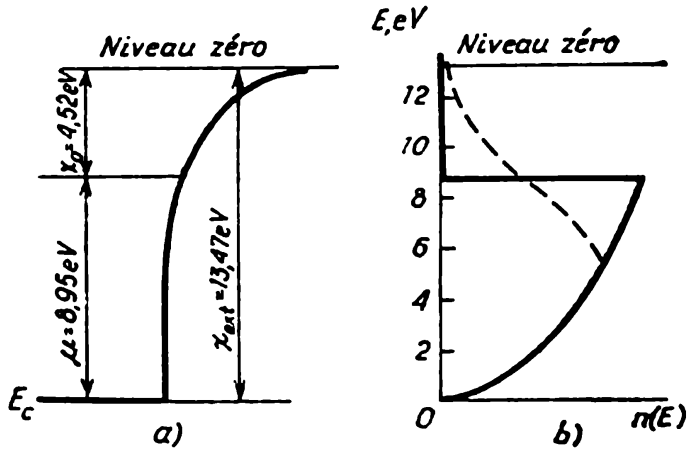

Fig. 10.4. Diagramme énergétique du tungstène (*a*) et courbes de répartition des électrons en fonction de l'énergie $n(E)$ à $T = 0$ K (courbe en traits pleins) et à une haute température (courbe en traits interrompus) (*b*).

la température s'élève, la « queue » de la courbe de distribution passe au-delà du niveau zéro vide, ce qui témoigne de l'apparition d'un certain nombre d'électrons possédant une énergie cinétique qui dépasse la hauteur de la barrière de potentiel. De tels électrons sont capables de quitter le métal (« s'évaporer »). Aussi, un métal échauffé émet-il des électrons. Ce phénomène a reçu le nom d'*émission thermoélectronique*. Elle ne devient notable qu'à une haute température, lorsque le nombre d'électrons thermiquement excités, capables de quitter le métal, devient suffisamment grand.

En plaçant au voisinage d'un métal chauffé un conducteur et en créant entre ce conducteur et le métal un champ capable de collecter les électrons, on peut obtenir un *courant thermoélectronique*.

Calculons la densité de ce courant en supposant que tous les électrons émis reviennent, par le circuit extérieur, sur la surface de sorte que l'électrode émettant les électrons reste dans son ensemble électriquement neutre. A cet effet il faut examiner de nouveau la distribution des électrons suivant les vitesses à l'intérieur du métal. Pour obtenir le nombre d'électrons par unité de volume de métal dont les composantes de la vitesse sont comprises dans les intervalles de v_x à $v_x + dv_x$, de v_y à $v_y + dv_y$ et de v_z à $v_z + dv_z$, il convient de multiplier le nombre d'états de l'électron dans ces intervalles par le degré de leur occupation. Comme il a été montré au § 5.2, les valeurs permises du vecteur d'onde **k** de l'électron sont distribuées dans l'espace des **k** uniformément de sorte qu'à chaque valeur de **k** correspond un volume de $8\pi^3$ (pour $V = 1$). Si nous utilisons la relation $m_n \mathbf{v} = \hbar \mathbf{k}$, qui est approximativement vérifiée pour les électrons de la bande de conduction, nous verrons que les valeurs permises du vecteur vitesse de l'électron sont uniformément distribuées dans l'espace des vitesses avec une densité $m_n^3/(8\pi^3\hbar^3) = m_n^3/h$. La densité d'états, compte tenu du spin des électrons, a une valeur deux fois plus grande $2m_n^3/h^3$. La probabilité d'occupation des états par les électrons ne dépend, suivant le § 5.1, que de l'énergie des électrons: $f_{F-D} = 1/\{\exp[(E-\mu)/(kT)]+1\}$. Alors, le nombre des électrons dans l'intervalle de composantes des vitesses $dv_x dv_y dv_z$ sera égal à

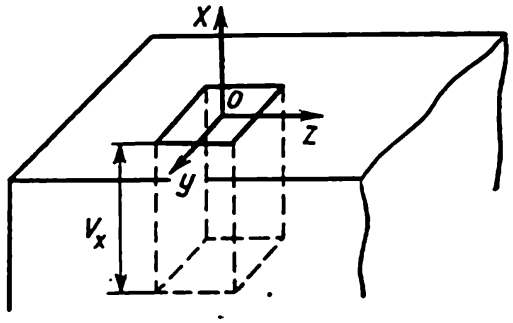

Fig. 10.5. Calcul de la densité de courant thermoélectrique.

$$n(v_x, v_y, v_z)\, dv_x\, dv_y\, dv_z = \frac{2m_n^3\, dv_x\, dv_y\, dv_z}{h^3\{\exp[(E-\mu)/(kT)]+1\}}. \qquad (10.2)$$

Séparons sur la surface de métal une unité de surface et, en l'utilisant comme base, construisons un parallélépipède rectangle d'arête latérale v_x (fig. 10.5). Le nombre des électrons contenus dans le parallélépipède, dont les composantes de la vitesse sont comprises dans les limites indiquées plus haut, a pour valeur

$$dN = v_x n(v_x, v_y, v_z)\, dv_x\, dv_y\, dv_z = \frac{2m_n^3 v_x\, dv_x\, dv_y\, dv_z}{h^3\{\exp[(E-\mu)/(kT)]+1\}}. \qquad (10.3)$$

Tous ces électrons atteignent en un intervalle de temps unité la surface du métal. Les électrons qui peuvent quitter le métal sont seulement ceux dont l'énergie cinétique dans le sens des x, égale à $m_n v_x^2/2$, sera non inférieure à la hauteur de la barrière χ_{ext}, c'est-à-dire ceux dont la vitesse dans le sens de l'axe des x ne sera pas inférieure à $v_{x0} = \sqrt{2\chi_{ext}/m_n}$.

En intégrant (10.3) par rapport à v_y et v_z dans les limites de $-\infty$ à $+\infty$ (les limites entre lesquelles peuvent théoriquement varier v_y et v_z) et par rapport à v_x dans les limites de v_{x0} à ∞, on obtient le nombre des électrons qui quittent chaque seconde une surface unité d'un métal chauffé :

$$N = \frac{2m_n^3}{h^3} \int_{v_{x0}}^{\infty} \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} \frac{v_x \, dv_x \, dv_y \, dv_z}{\exp[(E-\mu)/(kT)]+1}. \tag{10.4}$$

Le calcul de cette intégrale pour le cas généralement réalisé en pratique, lorsque $\chi_0 \gg kT$, conduit au résultat suivant :

$$N = 2\left(\frac{2\pi m_n k^2}{h^3}\right) T^2 \exp\left(-\frac{\chi_{ext}-\mu}{kT}\right). \tag{10.5}$$

Si l'on multiplie N par la charge de l'électron e, on obtient la densité de courant :

$$i = AT^2 \exp\left(-\frac{\chi_{ext}-\mu}{kT}\right) = AT^2 \exp\left(-\frac{\chi_0}{kT}\right), \tag{10.6}$$

où

$$A = 4\pi e m_n k^2/h^3. \tag{10.7}$$

La relation (10.6) exprime une *loi* dite *de Richardson-Dushmann*, dans laquelle A est appelé *constante de Richardson*. En prenant le logarithme de la relation (10.6), on obtient

$$\ln \frac{i}{T^2} = \ln A - \frac{\chi_0}{k}\frac{1}{T}. \tag{10.8}$$

Le graphique de cette fonction est donné sur la figure 10.6. C'est une droite dont l'inclinaison permet de déterminer le travail de sortie thermodynamique χ_0. En extrapolant cette droite jusqu'à l'intersection avec l'axe d'ordonnées, on peut trouver la constante de Richardson A. Pour $m_n = m_e$ la constante A doit être égale à $1{,}2 \times 10^6$ A/(m²·K²).

Ainsi qu'il ressort de la formule (10.6), la densité de courant thermoélectronique se détermine par la température de la surface émettrice et par le travail de sortie. Comme ces deux grandeurs interviennent dans l'exposant de l'exponentielle, la dépendance du courant en est très forte. C'est ainsi que si la température d'une cathode en tungstène est portée de 1000 à 2500 K, la densité de courant d'émission augmente d'environ 16 ordres ; le recouvrement d'une cathode en tungstène par une couche monoatomique de césium, qui diminue le travail d'extraction de 4,52 à 1,36 eV, a pour effet de provoquer une augmentation de la densité de courant d'environ 14 ordres de grandeur. C'est pourquoi de nos jours les cathodes en

métaux purs ne sont pas pratiquement utilisées (sauf des cathodes construites pour des cas spéciaux).

La relation (10.7) montre que la constante de Richardson doit être indépendante de la nature du métal. Mais en réalité elle est différente pour les divers métaux, ce qui est lié en partie au différent degré de transparence de leurs barrières de potentiel aux ondes électroniques qui se réfractent à la sortie des électrons du

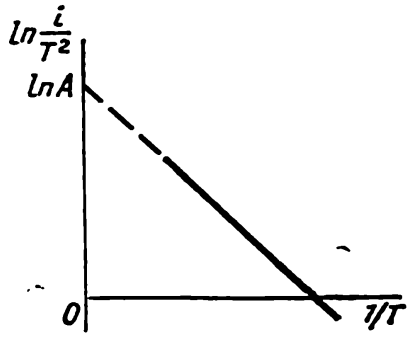

Fig. 10.6. Variation de la densité du courant de thermo-émission avec la température.

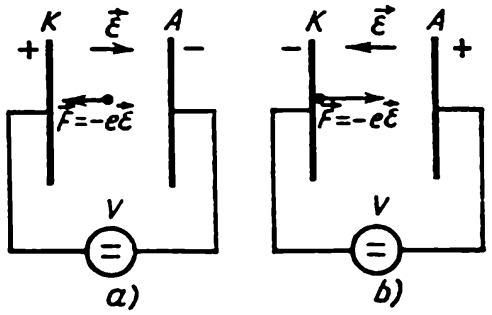

Fig. 10.7. Sens du champ extérieur $\mathcal{E}$ et de la force $\mathbf{F}$ agissant sur l'électron pour une tension d'anode retardatrice (*a*) et accélératrice (*b*).

métal dans le vide, mais se détermine essentiellement par les particularités de la fonction de densité d'états des métaux concrets au voisinage du niveau zéro.

Si l'on crée entre l'émetteur (la cathode thermo-électronique K; fig. 10.7, *a*) et le collecteur (l'anode A) une différence de potentiel V qui s'oppose au mouvement des électrons vers le collecteur, ce dernier ne pourra être atteint que par des électrons qui ont quitté l'émetteur avec une énergie cinétique non inférieure à $-eV$ ($V < 0$). Pour cela, leur énergie dans l'émetteur doit être non inférieure à $\chi_{ext} - eV$. En remplaçant dans l'expression (10.6) χ_{ext} par $\chi_{ext} - eV$, on obtient la relation suivante pour le courant circulant dans le circuit:

$$I = AST^2 \exp\left(-\frac{\chi_{ext} - eV - \mu}{kT}\right) = I_0 \exp\frac{eV}{kT}, \qquad (10.9)$$

où S est l'aire de la surface émettant les électrons; I_0 le courant d'émission pour $V = 0$. En prenant le logarithme de cette expression, on trouve

$$\ln I = \ln I_0 + eV/(kT). \qquad (10.10)$$

La figure 10.8, *a* montre la courbe de variation de $\ln I$ en fonction de V. Pour $V < 0$ c'est une droite qui coupe sur l'axe d'ordonnées ($V = 0$) un tronçon $\ln I_0$.

Lorsque le collecteur est porté à un potentiel positif (fig. 10.7, *b*), tous les électrons qui quittent l'émetteur atteignent le collecteur. Ceci étant, le courant dans le circuit ne doit pas varier tout en restant égal au courant de saturation I_0 (courbe en traits interrompus de la figure 10.8,*a*).

Il faut noter qu'une telle caractéristique courant-tension ne s'observe que dans le cas des valeurs relativement faibles des densités de courant d'émission et des potentiels positifs élevés du collecteur lorsqu'au voisinage de la surface émettrice il ne se forme pas une charge d'espace tant soit peu considérable des électrons qui n'ont pas

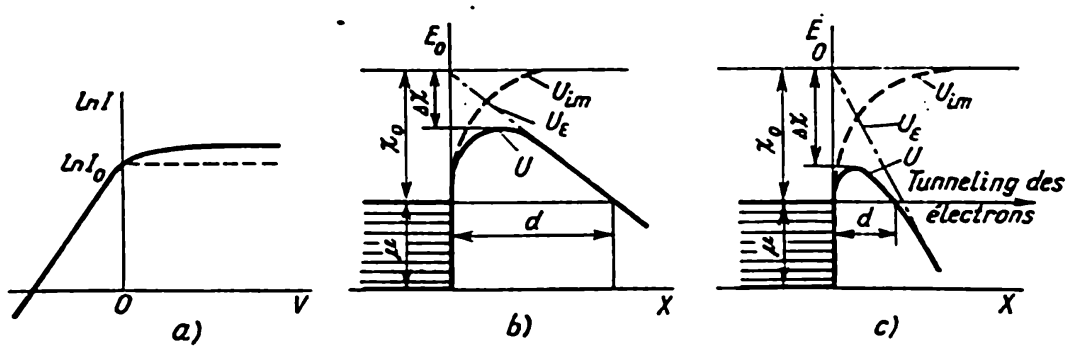

Fig. 10.8. Variation du courant en fonction de la tension d'anode (*a*) et influence d'un champ extérieur sur la hauteur et la forme de la barrière de potentiel à la surface de séparation métal-vide (*b*, *c*).

eu le temps pour atteindre le collecteur. En régime de charge d'espace la caractéristique courant-tension est déjà décrite par l'*équation de Child-Langmuir* suivant laquelle $I \sim V^{3/2}$.

Effet Schottky. La conclusion que pour $V > 0$ le courant I est indépendant de V n'est pas tout à fait exacte. Un champ accélérateur, régnant près de la surface émettrice et agissant sur l'électron avec une force $F = -e\mathscr{E}$ (fig. 10.7, *b*), effectue sur un parcours x un travail $Fx = -e\mathscr{E}x$ et diminue par là même l'énergie potentielle de l'électron d'une quantité égale à $|U_{\mathscr{E}}| = e\mathscr{E}x$. La figure 10.8, *b* montre la variation en fonction de x de l'énergie potentielle d'un électron dans un champ extérieur $U_{\mathscr{E}}$ (la droite en traits mixtes), dans le champ de forces d'image électrique U_{im} (la courbe en traits interrompus) et de l'énergie potentielle résultante U de l'électron (la courbe en traits pleins). Comme il est visible sur la figure 10.8, *b* le champ accélérateur, régnant près de la surface émettrice, *abaisse* la barrière de potentiel de $\Delta\chi$. Le calcul montre que pour des champs pas trop élevés

$$\Delta\chi = \sqrt{e^3\mathscr{E}/(4\pi\varepsilon_0)} \sim \sqrt{\mathscr{E}}. \qquad (10.11)$$

L'abaissement de la barrière de potentiel sous l'action d'un champ extérieur est connu sous le nom d'*effet Schottky*. Il a pour ré-

sultat que lorsque le potentiel positif du collecteur augmente, le courant d'émission ne reste pas constant (I_0) mais accuse une petite croissance (la courbe en traits pleins de la figure 10.8, *a*).

Emission froide d'électrons. Un champ accélérateur extérieur provoque non seulement l'abaissement de la barrière de potentiel mais aussi une diminution de son épaisseur d (fig. 10.8, *c*), ce qui rend une telle barrière, lorsque l'intensité de champ est très élevée ($\sim 10^9$ V/m), suffisamment transparente pour le passage des électrons par effet tunnel et pour leur sortie du solide. Ce phénomène est appelé *émission électronique froide* ou encore émission (électronique) par champ électrique. La densité de courant dans l'émission électronique par champ électrique croît brusquement avec l'augmentation de l'intensité du champ accélérateur:

$$i = C\mathscr{E}^2 \exp(-\alpha/\mathscr{E}), \tag{10.12}$$

où C et α sont des constantes caractéristiques de la barrière de potentiel.

§ 10.3. Différence de potentiel de contact

Considérons les processus qui se déroulent lorsque se rapprochent et se mettent en contact deux conducteurs électroniques, deux métaux par exemple, dont les diagrammes des énergies sont montrés sur

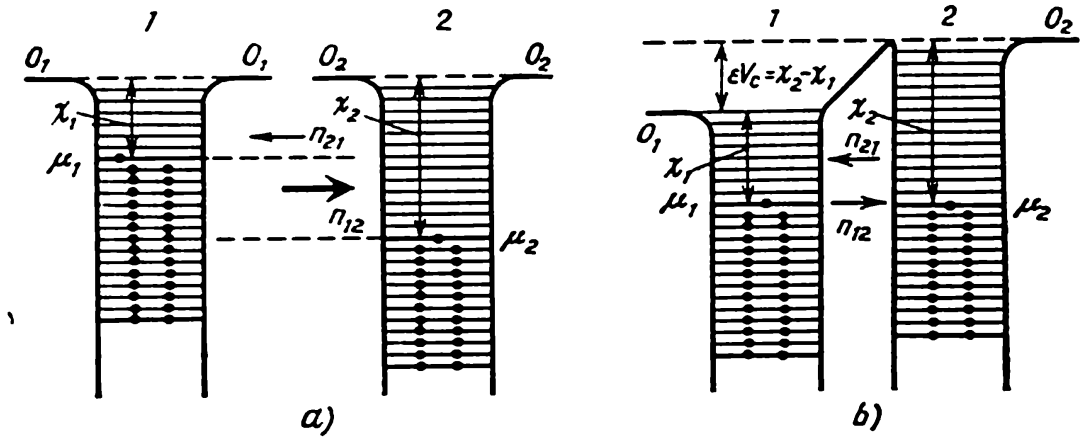

Fig. 10.9. Apparition d'une différence de potentiel de contact entre deux métaux.

la figure 10.9, *a*. Le gaz électronique de ces conducteurs se caractérise par les potentiels chimiques μ_1 et μ_2 et par les travaux d'extraction χ_1 et χ_2. Mettons ces conducteurs en contact, en les rapprochant l'un de l'autre à une distance à laquelle un échange effectif d'électrons devient possible. En examinant la figure 10.9, *a* qui montre comment sont occupés par les électrons les niveaux des métaux à $T = 0$ K, on voit que dans la bande de conduction du conducteur *2* sont occupés tous les états jusqu'au niveau de Fermi μ_2 et qu'en regard de ces

états se situent des niveaux occupés de la bande de conduction du conducteur *1*. Ceci étant, les électrons du conducteur *2* ne peuvent pas passer dans le conducteur *1*. A une température différente de 0 K, les électrons du conducteur *2* thermiquement excités à des niveaux situés plus haut que μ_1 peuvent passer dans le conducteur *1*, mais à des températures ordinaires le nombre de tels électrons n'est pas élevé de sorte que leur flux n_{21} est faible.

Il n'en est plus de même pour le conducteur *1*. Sa bande de conduction est, elle aussi, remplie jusqu'au niveau de Fermi μ_1, mais son travail d'extraction étant plus faible que celui du conducteur *2* ($\chi_1 < \chi_2$), en regard des états occupés de ce conducteur, situés plus haut que le niveau de Fermi μ_2, sont placés les niveaux libres de la bande de conduction du conducteur *2*. Aussi, lorsque les conducteurs sont mis en contact (même à $T = 0$ K), les électrons passent-ils des niveaux occupés du conducteur *1* aux niveaux libres du conducteur *2*, en constituant un flux n_{12} qui est au début beaucoup plus intense que le flux n_{21}.

Considérons maintenant comment s'établit un équilibre entre les conducteurs. Le conducteur *1*, qui perd des électrons, se charge positivement, le conducteur *2*, qui acquiert des électrons en excès, se charge négativement. L'apparition de ces charges a pour effet de provoquer un déplacement, l'un par rapport à l'autre, des niveaux énergétiques des deux conducteurs : dans le conducteur *1*, chargé positivement, tous les niveaux s'abaissent et dans le conducteur *2*, chargé négativement, tous les niveaux montent par rapport à leurs positions à l'état non chargé de ces conducteurs (fig. 10.9, *b*). Cela est facile à comprendre en partant des considérations bien simples suivantes. Pour faire passer un électron par exemple du niveau zéro O_1 du métal *1* au niveau zéro O_2 du métal *2* soumis à un potentiel négatif V par rapport au métal *1*, il faut dépenser un travail numériquement égal à eV. Ce travail se transforme en énergie potentielle de l'électron. Aussi, l'énergie potentielle de l'électron situé au niveau zéro d'un conducteur chargé négativement est de $\Delta U = eV$ plus grande que l'énergie potentielle d'un électron occupant le niveau zéro d'un conducteur chargé positivement. Mais cela signifie que le niveau zéro du conducteur *2* se situe de $\Delta U = eV$ plus haut que le niveau zéro du conducteur chargé positivement.

Un tel déplacement est subi par tous les niveaux énergétiques des conducteurs *1* et *2*, y compris les niveaux de Fermi μ_1 et μ_2.

Dès que le niveau de Fermi du conducteur *1* (μ_1) en descente continue et le niveau de Fermi du conducteur *2* (μ_2) en montée continue se trouvent à une même hauteur, la cause qui détermine le passage privilégié des électrons du conducteur *1* dans le conducteur *2* disparaît parce qu'en regard des niveaux occupés du conducteur *1* sont placés maintenant des niveaux occupés, avec le même degré, du conducteur *2*. Il s'établit entre les conducteurs un équilibre au-

quel correspond une différence de potentiel d'équilibre entre eux

$$V_c = \frac{1}{e}(\chi_2 - \chi_1), \tag{10.13}$$

que l'on appelle *différence de potentiel de contact*. Comme le montre la relation (10.13), V_c se détermine par la différence des travaux d'extraction des électrons des conducteurs en contact. La valeur de V_c peut atteindre quelques volts. L'examen de la figure 10.9, *b* montre que la différence de potentiel de contact crée pour les électrons passant dans le conducteur dont le travail d'extraction est plus grand, une barrière de potentiel haute de eV_c.

Le résultat obtenu est valable quel que soit le mode d'échange d'électrons entre deux matériaux, y compris par l'émission thermoélectronique dans le vide, à travers un circuit extérieur, etc.

La différence de potentiel de contact peut jouer un grand rôle dans le fonctionnement des appareils électroniques à vide dont les électrodes (la cathode, l'anode, les grilles et autres) sont faites, en règle générale, en métaux différents. La différence de potentiel de contact qui prend naissance entre de telles électrodes s'ajoute à la différence de potentiel extérieure et influe directement sur la caractéristique courant-tension des appareils. La variation de la différence de potentiel de contact peut provoquer un fonctionnement instable de ces appareils. Les phénomènes intervenant aux contacts sont également à la base de fonctionnement de nombreux appareils à semiconducteurs et d'autres appareils et dispositifs à corps solide.

Evaluons maintenant le nombre d'électrons passant d'un métal dans un autre lorsqu'il se crée une différence de potentiel de contact V_c. Nous supposerons qu'entre les métaux mis en contact il subsiste un jeu large de d et que les charges se concentrent dans des couches étroites près des surfaces. Alors, la valeur de la charge Q sur chacune des surfaces, nécessaire pour produire une différence de potentiel V_c peut être déterminée au moyen de la formule valable pour un condensateur plan $Q = CV_c$, où C est la capacité du condensateur. Dans le cas considéré $C = \varepsilon_0 S/d$ (ici S est l'aire des surfaces en contact, ε_0 la permittivité du vide). On en tire

$$Q = S\varepsilon_0 V_c/d.$$

Comme la distance d entre les métaux ne peut pas être inférieure au paramètre du réseau $a \simeq 0{,}3$ nm, la densité de charge maximale sur les surfaces en contact, pour une différence de potentiel de contact $V_c \approx 1$ V, est égale à $Q/S \simeq 3 \cdot 10^{-2}$ C/m^2. Cela correspond à $\simeq 2 \cdot 10^{17}$ électrons par 1 m^2. 1 m^2 de couche monoatomique de métal contient $\sim 5 \cdot 10^{18}$ atomes. En supposant que chacun de ces atomes cède au gaz électronique un électron de valence, on voit que la densité superficielle de gaz électronique $n_s \simeq 5 \cdot 10^{18}$ m^{-2}. Ainsi pour créer une différence de potentiel de contact il a fallu dans notre exemple

faire passer seulement $\simeq 4$ % d'électrons contenus dans la couche monoatomique près de la surface des métaux.

Une variation si petite de la concentration du gaz électronique dans la couche de contact d'une part et une faible épaisseur de la barrière de potentiel à la surface de séparation des métaux qui assure la transparence de cette barrière pour le passage des électrons par effet tunnel d'autre part, ne peuvent pas provoquer une variation tant soit peu notable de la conductivité de cette couche par rapport à celle du métal en profondeur. Le courant électrique passe par le contact de deux métaux avec la même facilité qu'à travers les métaux eux-mêmes.

§ 10.4. Contact entre un métal et un semiconducteur

Contact de blocage. Considérons le contact entre un métal et un semiconducteur. Soit un métal M ayant le travail d'extraction χ_{0m} mis en contact avec un semiconducteur électronique Π ayant le travail d'extraction χ_{0s} (fig. 10.10). Pour $\chi_{0m} > \chi_{0s}$ les électrons

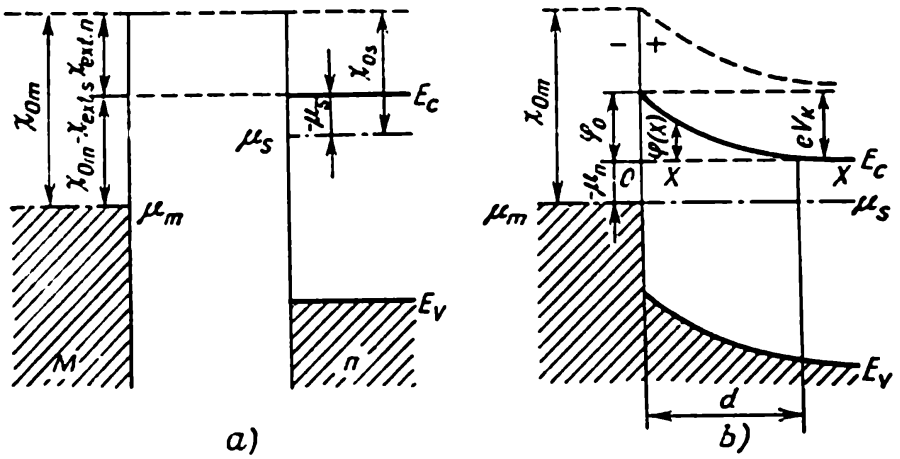

Fig. 10.10. Schéma d'un contact métal-semiconducteur.

du semiconducteur passeront dans le métal jusqu'à l'instant où les potentiels chimiques μ_m et μ_s s'égalisent et un équilibre s'établit. Entre le métal et le semiconducteur il apparaît une différence de potentiel de contact V_k qui est approximativement du même ordre de grandeur que dans le cas du contact de deux métaux (de l'ordre de 1 V). Pour créer une telle différence de potentiel il faut que le nombre d'électrons passant du semiconducteur dans le métal soit à peu près le même que dans le cas du contact de deux métaux.

Si le paramètre du réseau du semiconducteur $a \approx 0{,}5$ nm (germanium) et la concentration du gaz électronique $n \approx 10^{21}$ m^{-3}, 1 m^2 de sa surface contient dans la couche monoatomique $n_s \approx 10^{14}$ électrons. Aussi le passage de $\Delta n \approx 2 \cdot 10^{17}$ électrons doit être lié à la « mise à nu » de $\sim 10^3$ couches atomiques du semiconducteur.

Ainsi, dans le cas du contact entre des métaux et un semiconducteur l'égalisation des potentiels chimiques se produit par le passage sur la surface de contact du métal des électrons à partir de la couche frontière du semiconducteur d'épaisseur considérable d (fig. 10.11) et toute la différence de potentiel de contact chute dans le semiconducteur et non dans le jeu entre le semiconducteur et le métal. Les atomes d'impureté ionisés, qui restent dans cette couche, constituent une charge d'espace positive immobile. Comme cette couche ne contient pratiquement pas d'électrons libres, on l'appelle couche *appauvrie* et le contact métal-couche appauvrie de semiconducteur qui a une grande résistance est appelé *contact de blocage*.

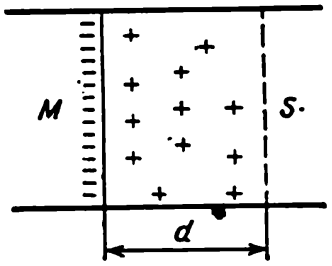

Fig. 10.11. Répartition des charges dans un contact de blocage métal-semiconducteur.

Influence du champ de contact sur les niveaux énergétiques d'un semiconducteur. La différence de potentiel de contact V_k qui prend naissance entre un métal et un semiconducteur est formée suivant toute l'épaisseur d de la couche appauvrie (fig. 10.11). Si l'on prend pour la constante de réseau $a \approx 0{,}5$ nm et pour le nombre de couches de contact « dénudées » du semiconducteur $\sim 10^3$, on obtient pour l'épaisseur de cette couche $d \approx 0{,}5 \cdot 10^3$ nm $= 5 \cdot 10^{-7}$ m. Pour $V_k = 1$ V l'intensité du champ de contact $\mathscr{E}_k \approx V_k/d \approx 2 \cdot 10^6$ V/m. C'est au moins de trois ordres de grandeur inférieur à l'intensité du champ interne du cristal, qui détermine le spectre énergétique du semiconducteur. Cela signifie que le champ de contact ne peut pas avoir une influence appréciable sur la structure de ce spectre (la largeur de la bande interdite, l'énergie d'activation des impuretés, etc.). Son action se réduit seulement à l'incurvation de tous les niveaux énergétiques du semiconducteur. Expliquons ceci.

En l'absence de champ de contact, les niveaux énergétiques dans le métal et dans le semiconducteur sont représentés par des droites horizontales (v. fig. 10.10, *a*). On exprime ainsi le fait que l'énergie d'un électron occupant un niveau donné, par exemple le niveau inférieur de la bande de conduction, est la même dans tous les points du semiconducteur. Elle ne dépend pas des coordonnées des électrons. En présence de la différence de potentiel de contact, la situation est différente : dans la couche où est concentré le champ de contact, l'électron est soumis à une force qui tend à le chasser de la couche. Pour surmonter cette force il faut dépenser un travail qui se transforme en énergie potentielle de l'électron. C'est pourquoi, au fur et à mesure que l'électron se déplace à l'intérieur de la couche de charge d'espace, son énergie potentielle $\varphi(x)$ augmente pour atteindre sa valeur ma-

ximale $\varphi_0 = e \mid V_k \mid$ sur la frontière du semiconducteur (v. fig. 10.10, *b*). Cela signifie que le champ de contact provoque une courbure des bandes d'énergie du semiconducteur.

Pour déterminer la forme de la fonction $\varphi(x)$ utilisons l'équation de Poisson (8.51):

$$\frac{d^2V}{dx^2} = -\frac{\rho(x)}{\varepsilon\varepsilon_0},$$

où ε est la constante diélectrique du semiconducteur.

Il est raisonnable de passer du potentiel $V(x)$ à l'énergie potentielle de l'électron $\varphi(x) = -eV(x)$ et d'écrire l'équation de Poisson pour $\varphi(x)$:

$$\frac{d^2\varphi}{dx^2} = \frac{e}{\varepsilon\varepsilon_0}\rho(x). \tag{10.14}$$

Pour calculer la densité volumique $\rho(x)$ de la charge nous supposerons que dans la couche appauvrie tous les atomes donneurs N_d sont ionisés et que tous les électrons sont passés dans le métal. Alors $\rho = eN_d$. En portant la valeur de μ dans l'équation de Poisson, on obtient

$$\frac{d^2\varphi}{dx^2} = \frac{e^2}{\varepsilon\varepsilon_0}N_d. \tag{10.15}$$

Etant donné qu'à une distance $x \geqslant d$ le champ de contact n'existe pas dans le semiconducteur, les conditions aux limites pour cette équation peuvent s'écrire sous la forme

$$\varphi(d) = 0; \quad \left(\frac{d\varphi}{dx}\right)_{x=d} = 0. \tag{10.16}$$

L'intégration de l'équation (10.15) avec les conditions aux limites (10.16) conduit au résultat suivant:

$$\varphi(x) = \frac{e^2N_d}{2\varepsilon\varepsilon_0}(d-x)^2. \tag{10.17}$$

La relation (10.17) montre que lorsque x augmente, le potentiel de champ de contact dans le semiconducteur décroît suivant une loi parabolique. Pour $x = 0$ on obtient $\varphi_0 = \chi_{0m} - \chi_{0s}$. En portant la valeur de φ_0 dans (10.17), on trouve l'épaisseur d de la couche appauvrie:

$$d = \sqrt{\frac{2\varepsilon_0\varepsilon\varphi_0}{e^2N_d}} = \sqrt{\frac{2\varepsilon\varepsilon_0\varphi_0}{e^2n_{n0}}}, \tag{10.18}$$

où n_{n0} est la concentration des électrons (des porteurs majoritaires) dans le semiconducteur de type n qui est égale à N_d.

Il résulte de (10.18) que l'épaisseur d de la couche appauvrie augmente en même temps que la différence de potentiel de contact

φ_0/e déterminée par la différence des travaux d'extraction et diminue lorsque la concentration des porteurs majoritaires augmente dans le semiconducteur. Les résultats de calcul de d au moyen de (10.18) dans l'hypothèse que $\varphi_0 = 1$ V et $\varepsilon = 10$ sont indiqués dans le tableau 10.1.

Tableau 10.1

n_{n0}, m^{-3}	d, m
10^{24}	$3 \cdot 10^{-8}$
10^{22}	$3 \cdot 10^{-7}$
10^{20}	$3 \cdot 10^{-6}$

Si le travail de sortie d'un semiconducteur électronique est plus grand que le travail de sortie du métal, les électrons passent du métal dans le semiconducteur en constituant dans sa couche de contact une charge d'espace négative. A mesure que l'électron se déplace vers la surface du semiconducteur, son énergie $\varphi(x)$ n'augmente pas dans ce cas mais au contraire diminue de sorte que les bandes d'énergie s'incurvent dans le sens opposé (fig. 10.12). La concentration des charges mobiles dans la couche de contact du semiconducteur devient dans ce cas plus forte. C'est pourquoi une telle couche est dite *enrichie*. Près de la surface frontière entre un semiconducteur de type p et un métal peuvent se former aussi bien des couches d'appauvrissement que des couches d'enrichissement en trous.

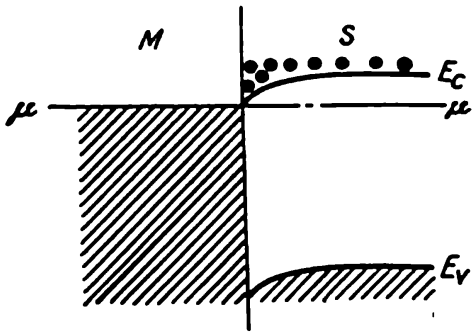

Fig. 10.12. Diagramme énergétique d'un contact non redresseur métal-semiconducteur de type n.

Redressement par un contact entre un semiconducteur et un métal. Une propriété remarquable du contact de blocage est une très forte variation de sa résistance en fonction du sens de la différence de potentiel extérieure appliquée au contact. Cette variation est si forte qu'elle conduit pratiquement à une conductibilité unidirectionnelle (*unipolaire*) du contact : le contact conduit facilement le courant dans le *sens* dit *direct* et beaucoup plus difficilement dans le sens opposé (inverse). C'est là que réside l'effet redresseur d'un contact entre un semiconducteur et un métal. Examinons-le de plus près.

La figure 10.13, *a* montre la structure de bande d'un contact entre un semiconducteur électronique et un métal à l'état d'équilibré.

Pour les électrons passant du métal dans le semiconducteur la barrière de potentiel est égale à la différence $\chi_{0m} - \chi_{ext.s}$; pour les électrons passant du semiconducteur dans le métal elle est égale à $\varphi_0 = e \mid V_k \mid$. A l'état d'équilibre, le flux d'électrons passant du métal dans le semiconducteur est égal au flux opposé d'électrons passant du semiconducteur dans le métal de sorte que le courant résultant à travers le contact est nul. Désignons par i_s la densité de courant correspondant aux densités d'équilibre des flux électroniques. Elle est égale à $\frac{1}{4}\langle v_n \rangle e n_{n0} e^{-\varphi_0/kT}$, où $\langle v_n \rangle$ est la vitesse moyenne de l'agitation thermique des électrons dans le semiconducteur.

Appliquons au contact une différence de potentiel extérieure dans le sens coïncidant avec la différence de potentiel de contact V_k, en chargeant le semiconducteur positivement par rapport au métal

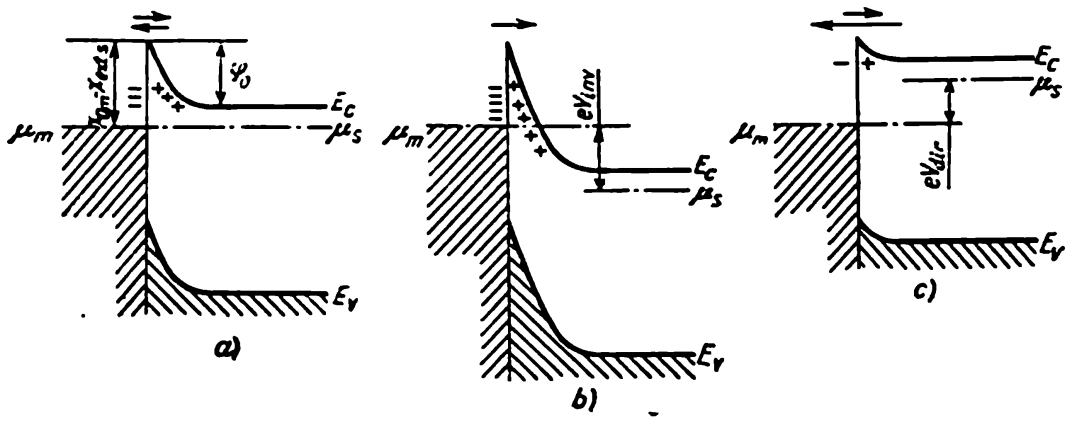

Fig. 10.13. Diagramme énergétique d'un contact redresseur métal-semiconducteur à l'équilibre (*a*) et en polarisation inverse V_{inv} (*b*) et direct V_{dir} (*c*) (les flèches indiquent les flux d'électrons).

(fig. 10.13, *b*); un tel sens est dit *inverse*. Convenons de considérer que le signe de la polarisation extérieure V appliquée dans le sens inverse est négatif. Etant donné que la résistance du contact est généralement de plusieurs ordres de grandeur supérieure à celle du reste du circuit, pratiquement toute la différence de potentiel V appliquée est concentrée sur le contact. Les niveaux d'énergie dans un semiconducteur chargé positivement sont déplacés de eV vers le bas par rapport à leurs positions initiales. Le niveau de Fermi μ_s descendra de la même distance. Sur la figure 10.13, *b* le niveau inférieur de la bande de conduction E_c et le sommet de la bande de valence E_v sont construits compte tenu de cette nouvelle position du niveau de Fermi. Comme il est visible sur la figure, la différence de potentiel extérieure V appliquée dans le sens inverse provoque une augmentation de la barrière de potentiel pour les électrons passant du semiconducteur dans le métal jusqu'à

$$\varphi(0) = \varphi_0 - eV \text{ (pour } V < 0). \qquad (10.19)$$

Suivant (10.18) la couche appauvrie correspondant à une telle barrière a l'épaisseur

$$d=\sqrt{\frac{2\varepsilon\varepsilon_0(\varphi_0-eV)}{e^2n_{n0}}} \text{ (pour } V<0). \qquad (10.20)$$

Ainsi, un champ extérieur appliqué au contact dans le sens inverse élève la barrière de potentiel pour les électrons passant du semiconducteur dans le métal et augmente l'épaisseur de la couche appauvrie.

La situation est différente dans le cas où le champ est appliqué au contact dans le sens direct (fig. 10.13, *c*). Dans ce cas, dans le semiconducteur chargé négativement, tous les niveaux d'énergie, y compris le niveau de Fermi μ_s, sont déplacés vers le haut de eV, ce qui conduit à la réduction de eV de la barrière énergétique pour les électrons passant du semiconducteur dans le métal. La barrière devient égale à

$$\varphi(0)=\varphi_0-eV \text{ (pour } V>0). \qquad (10.21)$$

Respectivement, l'épaisseur de la couche de charge d'espace tombe jusqu'à

$$d=\sqrt{\frac{2\varepsilon\varepsilon_0(\varphi_0-eV)}{e^2n_{n0}}} \text{ (pour } V>0). \qquad (10.22)$$

La variation de la hauteur de la barrière de potentiel sous l'action d'une différence de potentiel extérieure trouble l'équilibre entre les flux d'électrons passant par le contact. Si l'on applique au contact une différence de potentiel extérieure V dans le sens inverse, la densité de courant i_{nm} deviendra $e^{e|V|/(kT)}$ fois plus petite parce qu'en vertu de la loi de Boltzmann le nombre d'électrons passant du semiconducteur dans le métal, capables de franchir une barrière de hauteur $\varphi_0+e|V|$ est $e^{e|V|/(kT)}$ fois plus petit que le nombre d'électrons capables de surmonter une barrière d'équilibre φ_0. C'est pourquoi la densité de courant d'électrons passant du semiconducteur dans le métal devient égale à

$$i_{nm}=i_s e^{eV/(kT)} \text{ (pour } V<0). \qquad (10.23)$$

Quant à la densité de courant opposé i_{mn}, elle reste égale à i_s parce que le champ extérieur est sans effet sur la hauteur de la barrière pour les électrons passant du métal dans le semiconducteur: elle reste toujours égale à la différence des travaux d'extraction $\chi_{0m}-\chi_{ext.s}$.

La densité résultante du courant circulant dans le sens inverse est égale à

$$i=i_s e^{eV/(kT)}-i_s=i_s(e^{eV/(kT)}-1) \text{ (pour } V<0). \qquad (10.24)$$

Ce courant est dirigé du semiconducteur vers le métal. Si la tension inverse V augmente, l'exponentielle $e^{eV/(kT)}$ tend rapidement vers zé-

ro et la densité de courant inverse vers une valeur limite égale à $-i_s$. La densité de courant i_s est appelée *densité de courant de saturation* et $I_s = Si_s$ est dit courant de saturation (S étant l'aire de la section droite du contact entre le métal et le semiconducteur).

Si la différence de potentiel extérieure V est appliquée dans le sens direct (fig. 10.13, *c*), la barrière de potentiel pour les électrons passant du semiconducteur dans le métal s'abaisse de eV de sorte que la densité de courant i_{nm} constitué par ces électrons augmente de $e^{eV/(kT)}$ fois par rapport à i_s :

$$i_{nm} = i_s e^{eV/(kT)} \text{ (pour } V > 0).$$

Quant à la densité de courant i_{mn} elle est toujours égale à i_s. Cela signifie que la densité de courant passant dans le sens direct (du métal dans le semiconducteur) a pour valeur

$$i = i_{nm} - i_{mn} = i_s (e^{eV/(kT)} - 1) \quad \text{pour } (V > 0). \qquad (10.25)$$

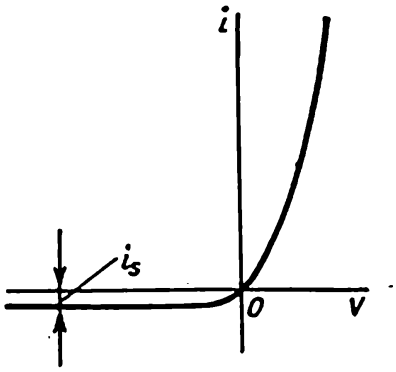

Fig. 10.14. Caractéristique courant-tension d'une diode Schottky.

Ce courant augmente exponentiellement lorsque V augmente. En réunissant (10.24) et (10.25), on obtient

$$i = i_s (e^{eV/(kT)} - 1). \qquad (10.26)$$

Dans ce cas $V > 0$ pour la polarisation directe et $V < 0$ pour la polarisation inverse.

La formule (10.26) est l'*équation de la caractéristique courant-tension* (CCT) du contact redresseur entre un métal et un semiconducteur. La forme d'une telle caractéristique est montrée sur la figure 10.14. L'examen de cette figure montre que le contact entre un métal et un semiconducteur possède réellement une propriété de redressement : il laisse passer le courant dans le sens direct et le bloque presque totalement dans le sens inverse.

Le rapport de l'intensité de courant passant dans le sens direct à l'intensité de courant circulant dans le sens inverse correspondant à une même différence de potentiel est appelé *coefficient de redressement*. Pour des contacts redresseurs de bonne qualité ce coefficient atteint des valeurs se chiffrant par des dizaines et même des centaines de milliers.

La barrière de potentiel qui prend naissance au contact redresseur semiconducteur-métal est souvent appelée *barrière de Schottky* et les diodes utilisant cette barrière portent le nom de *diodes Schottky*.

§ 10.5. Courants limités par la charge d'espace

Les contacts non redresseurs entre un métal et un semiconducteur dont la formation s'accompagne d'une création dans la région de contact du semiconducteur d'une couche enrichie en porteurs de courant majoritaires n'apportent aucune résistance additionnelle au passage du courant et s'appellent *contacts ohmiques*. Les contacts ohmiques

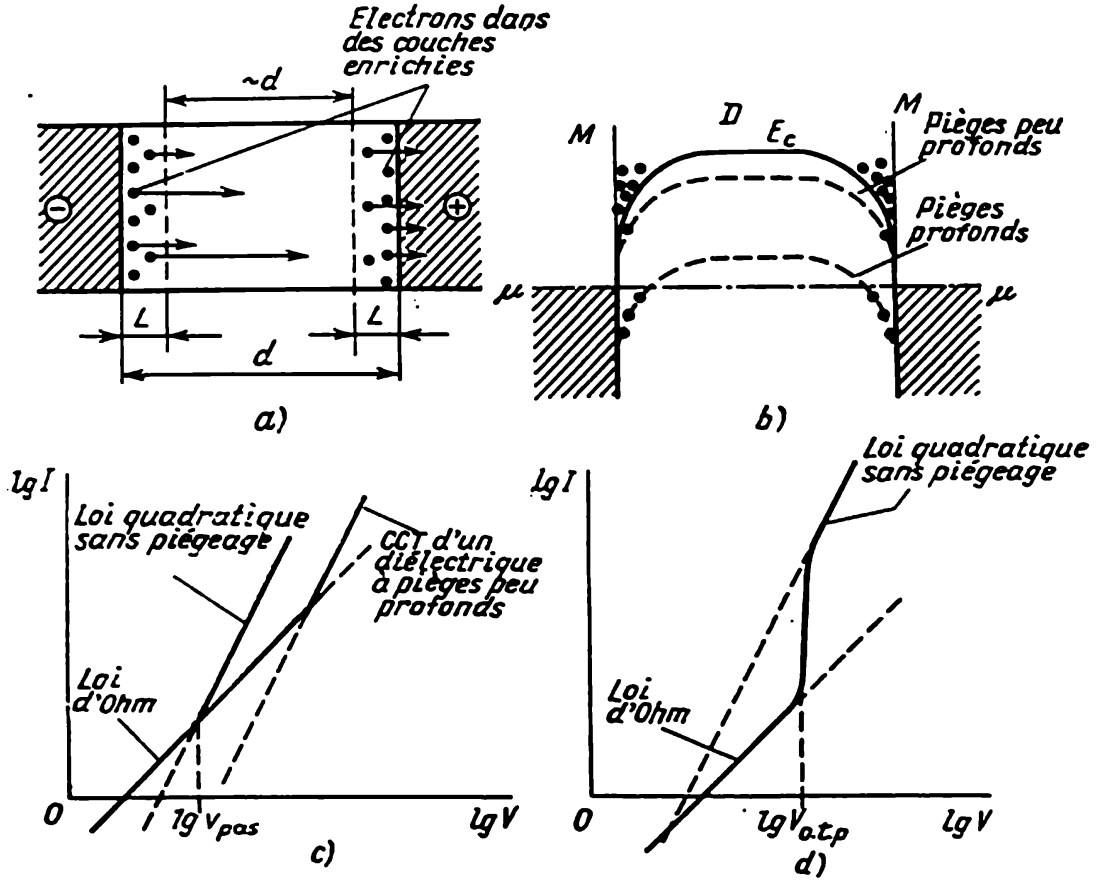

Fig. 10.15. Courants limités par une charge d'espace:
a — structure MDM ; *b* — diagramme énergétique d'une structure MDM à l'état d'équilibre ; *c* — courbe caractéristique courant-tension d'une structure MDM avec un diélectrique sans pièges et avec un diélectrique avec des pièges peu profonds ; *d* — caractéristique courant-tension d'une structure MDM avec un diélectrique contenant des pièges profonds.

sont nécessaires pour l'établissement des connexions entre les régions semiconductrices des appareils et les circuits électriques extérieurs. Dans la réalité, le passage à travers le contact du courant d'une telle polarité, lorsque les porteurs majoritaires se dirigent du contact vers le semiconducteur, a pour effet même une certaine diminution de la résistance du semiconducteur du fait que les porteurs majoritaires sont entraînés (injectés) de la région de contact dans la profondeur du semiconducteur. Cet effet est particulièrement marqué dans les structures métal-diélectrique-métal (MDM) où le courant est souvent

transporté précisément par des porteurs injectés et non pas par des porteurs intrinsèques peu nombreux.

Considérons le cas où l'épaisseur d du diélectrique est beaucoup plus grande que l'épaisseur L des couches enrichies qui apparaissent près des contacts métalliques (fig. 10.15, *a*, *b*).

Diélectrique parfait. Commençons notre étude par l'examen d'un « diélectrique parfait » ne contenant ni impuretés, ni centres capables de capter les électrons. La concentration des porteurs de charges intrinsèques dans un tel diélectrique est infime. Aussi, une tension extérieure V appliquée à une telle structure chute-t-elle pratiquement en totalité dans le diélectrique, à savoir dans celle de ses parties où n'entrent pas les couches d'enrichissement, et y produit un champ électrique (en plus du champ intérieur près des contacts). N'ayant pas de frontière nette, la couche enrichie du domaine cathodique se propage à l'intérieur du diélectrique en constituant une charge d'espace de densité variable ; les électrons de la couche enrichie près de l'anode passent dans le circuit extérieur. Le courant électrique qui en résulte est limité, de même que dans le cas d'une diode à vide, lorsque les densités du courant d'émission sont grandes, par la charge d'espace des porteurs distribués dans le diélectrique. C'est pourquoi on l'appelle *courant limité par la charge d'espace* (CLCE).

Evaluons l'ordre de grandeur du nombre d'électrons injectés de cette façon dans la couche diélectrique. La charge Q de la couche diélectrique est liée à la différence de potentiel appliquée V par la relation

$$Q = CV, \tag{10.27}$$

où C est la capacité du système. Si les charges additionnelles ne se formaient que près des frontières du film diélectrique, la capacité serait égale à

$$C = \varepsilon\varepsilon_0 S/d, \tag{10.28}$$

où S est la surface des contacts. D'un autre côté, si les électrons injectés dans le film de diélectrique y étaient distribués uniformément, il faudrait écrire dans (10.28) $d/2$ au lieu de d parce que la distance moyenne entre les « plus » et les « moins » serait dans ce cas deux fois plus petite. Alors

$$C = 2\varepsilon\varepsilon_0 S/d. \tag{10.29}$$

On peut montrer que la distribution réelle des électrons dans le diélectrique est intermédiaire entre les deux cas extrêmes considérés et que les valeurs de la capacité sont comprises entre $\varepsilon\varepsilon_0 S/d$ et $2\varepsilon\varepsilon_0 S/d$. Pour évaluer l'ordre de grandeur de C on peut se servir de n'importe laquelle de ces valeurs, par exemple (10.29). Alors

$$Q = 2\varepsilon\varepsilon_0 SV/d. \tag{10.30}$$

Cette charge est liée à la concentration n des électrons dans le diélectrique par la relation

$$Q = eS \int_0^d n(x)\,\mathrm{d}x.$$

La concentration moyenne des électrons dans le diélectrique a pour expression

$$\bar{n} = \frac{1}{d} \int_0^d n(x)\,\mathrm{d}x = \frac{Q}{eSd} = \frac{2\varepsilon\varepsilon_0}{e} \frac{V}{d^2}, \qquad (10.31)$$

et la conductance du diélectrique

$$G \approx \bar{\sigma} S/d = e\bar{n} u_n S/d,$$

où $\bar{\sigma}$ est la conductibilité moyenne produite dans le diélectrique; u_n la mobilité des électrons dans le diélectrique. En y portant la valeur de $\bar{n}$ donné par (10.31), on obtient

$$G \approx 2\varepsilon\varepsilon_0 u_n SV/d^3. \qquad (10.32)$$

Le diélectrique sera parcouru par un courant

$$I = GV \approx 2\varepsilon\varepsilon_0 u_n SV^2/d^3. \qquad (10.33)$$

En réalité, les électrons sont répartis suivant l'épaisseur du film inégalement: leur concentration près de la cathode est plus forte, mais l'intensité de champ électrique y est plus faible. Cela assure une densité de courant constante dans toute section du film.

Un calcul plus précis, fondé sur la résolution de l'équation de Poisson et de l'équation de continuité de courant, donne au lieu de (10.33) l'expression suivante:

$$I = {}^9/_8 \varepsilon\varepsilon_0 u_n SV^2/d^3. \qquad (10.34)$$

C'est une *loi* dite *quadratique sans piégeage* suivant laquelle *le courant limité par la charge d'espace est proportionnel au carré de la tension appliquée à la structure MDM*. La caractéristique courant-tension d'une telle structure est montrée en coordonnées logarithmiques sur la figure 10.15, *c*.

Diélectrique imparfait. Si avant l'application d'une tension de polarisation extérieure le film contenait des porteurs de charges libres de concentration n_0 (un diélectrique imparfait ou un semiconducteur) le courant circulant à travers le film sera la somme du courant transporté par ces porteurs et du courant de porteurs ayant passé dans le film de la couche enrichie. Lorsque les polarisations sont faibles, la concentration de ces derniers est faible par rapport à n_0 de sorte que le courant résultant est proportionnel à la tension de polarisation appliquée. Dans le cas de fortes tensions de polarisation, lors-

que le nombre de porteurs injectés devient supérieur au nombre de porteurs en équilibre, le courant devient une fonction quadratique de la tension de polarisation (fig. 10.15, *c*). La tension V_{pas} pour laquelle il se produit le passage de la loi d'Ohm à la dépendance quadratique de I par rapport à V, est d'autant plus élevée que la concentration initiale n_0 de porteurs est plus forte.

Diélectrique à pièges. Dans la bande interdite du diélectrique il existe souvent des niveaux localisés, créés par des atomes d'impureté ou par des défauts du réseau, qui sont capables de capturer les électrons injectés dans le film. Ces pièges peuvent se situer à une profondeur différente dans la bande interdite. Les pièges sont dits peu profonds s'ils se situent si près de la bande de conduction que loin du contact ils sont nettement beaucoup plus hauts que le niveau de Fermi et faiblement occupés par les électrons (fig. 10.15, *b*). Entre ces pièges et la bande de conduction il s'effectue constamment un échange d'électrons. Si l'on désigne par δ le rapport du temps de séjour des électrons dans la bande de conduction au cours duquel ils peuvent participer à la formation du courant au temps total de leur séjour dans le diélectrique, l'intensité du courant dans le diélectrique en présence de pièges peut être déterminée en multipliant (10.34) par δ:

$$I = {}^9/_8 \delta \varepsilon \varepsilon_0 u_n S V^2/d^3. \tag{10.35}$$

Comme $\delta < 1$, toutes choses étant égales par ailleurs, le courant dans un diélectrique avec pièges sera moins intense que dans un diélectrique sans pièges bien que toujours $I \sim V^2$ (fig. 10.15, *c*).

Les pièges profonds sont complètement occupés près du contact avec le métal, mais ils peuvent être en partie libres loin du contact (fig. 10.15, *b*). Les électrons injectés dans le diélectrique, lorsqu'il est soumis à une tension de polarisation extérieure, sont capturés par de tels pièges et bien retenus. Ils passent très rarement dans la bande de conduction. Aussi, tant que les porteurs injectés n'occupent pas ces pièges, le courant dans le diélectrique se détermine principalement par ses propres porteurs. Après une occupation totale des pièges pour une certaine tension $V_{o.t.p}$, les porteurs injectés sont dirigés seulement dans la bande de conduction et le courant accuse une brusque croissance. Lorsque la tension $V_{o.t.p}$ est dépassée de façon appréciable, la concentration de porteurs dans la bande de conduction devient beaucoup plus forte que leur concentration dans les pièges et la caractéristique courant-tension de la structure se décrit par une loi quadratique sans piégeage (fig. 10.15, *d*).

Injection double. Jusqu'ici nous n'avons examiné que l'injection d'électrons dans un diélectrique, c'est-à-dire considéré que les deux contacts étaient passants pour les électrons et bloquants pour les trous ou les deux contacts étaient passants pour les trous et bloquants pour les électrons.

Le mécanisme physique de passage du courant à travers le film diélectrique change notablement lorsqu'un de ses contacts est passant pour les électrons et l'autre l'est pour les trous. Le diagramme des énergies d'une telle structure est représenté sur la figure 10.16, *a*. Une telle structure peut être réalisée en principe si l'on prend pour le contact de droite un métal dont le travail d'extraction est inférieur et pour le contact de gauche un métal dont le travail d'extraction est

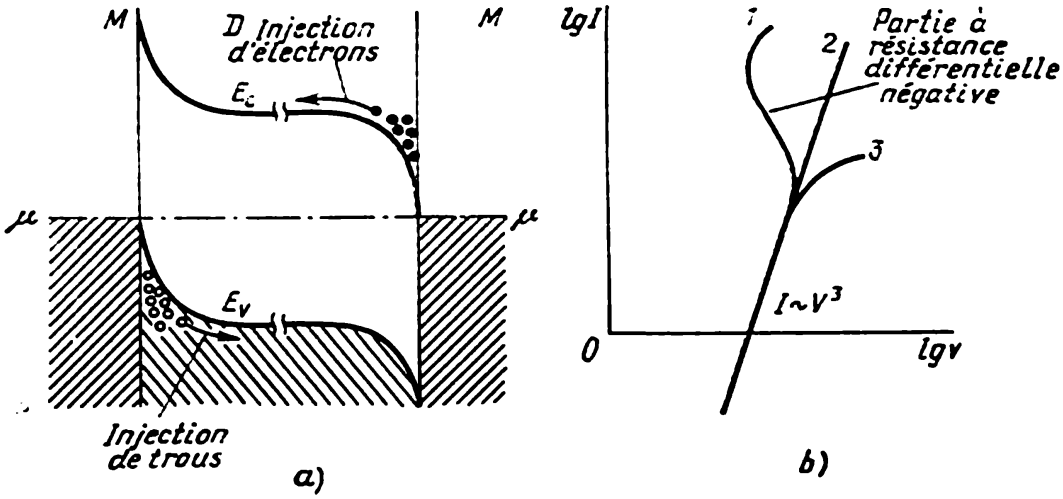

Fig. 10.16. Double injection:
a — diagramme énergétique d'une structure MDM dans laquelle le contact de droite sert à l'injection d'électrons et le contact de gauche à l'injection de trous; *b* — caractéristique courant-tension d'une structure MDM dans le cas d'une double injection.

supérieur au travail d'extraction du diélectrique. Dans ce cas la région de contact de droite sera enrichie en électrons et la partie de gauche le sera en trous.

La caractéristique courant-tension d'une telle structure est asymétrique. Si le contact de droite est relié au pôle « + » de la source de tension de polarisation et le contact de gauche au pôle « — », la résistance de la structure devient très grande. Au contraire, si le pôle « + » de la source est relié au contact de gauche et le pôle « — » au contact de droite, les électrons et les trous des régions de contact seront injectés dans la profondeur du film diélectrique en assurant une conductibilité appréciable. De même que dans le cas de l'injection dans le diélectrique de porteurs de charge d'un même signe (*injection monopolaire*), que nous avons examiné plus haut, la charge des porteurs injectés est limitée par la valeur de CV, où C est la capacité de la structure et V la tension de polarisation appliquée à la structure. Mais il est à noter que les charges d'espace apparaissent maintenant dans le film du fait que les concentrations des électrons et des trous injectés sont différentes. Aux concentrations elles-mêmes aucune limitation n'est imposée. C'est pourquoi dans le cas de l'*injec-*

tion double (*bipolaire*) le courant est plus intense, pour une même différence de potentiel, que dans l'injection monopolaire et, comme le montre le calcul, il se détermine par la relation suivante:

$$I = 2\varepsilon\varepsilon_0 u^2 S\tau V^3/d^5 \sim \tau V^3/d^5, \qquad (10.36)$$

où τ est la durée de vie des porteurs de charges.

Comme le montre (10.36), si la durée de vie des porteurs de charges ne varie pas avec le niveau d'injection, le courant à travers le film diélectrique est proportionnel à V^3 (courbe *2* de la figure 10.16, *b*). Dans les cas où la durée de vie des porteurs de charges dépend essentiellement du niveau de leur injection dans le film, la dépendance $I \sim V^3$ cesse d'être valable. Dans ce cas, si la durée de vie τ diminue lorsque la concentration n des porteurs injectés augmente, le courant varie avec la tension plus lentement que V^3 (courbe *3* de la figure 10.16, *b*) et, au contraire, si la durée τ augmente avec l'augmentation de n, le courant I croît plus vite que V^3. Dans ce dernier cas, si la durée de vie croît avec le courant suffisamment vite, la caractéristique courant-tension de la structure peut comporter une région à résistance différentielle négative et la caractéristique courant-tension elle-même prend la forme en S (courbe *1* de la figure 10.16, *b*).

§ 10.6. Force thermo-électromotrice

Les forces électromotrices de contact ne peuvent pas provoquer l'apparition d'un courant électrique dans un circuit fermé si tous ses

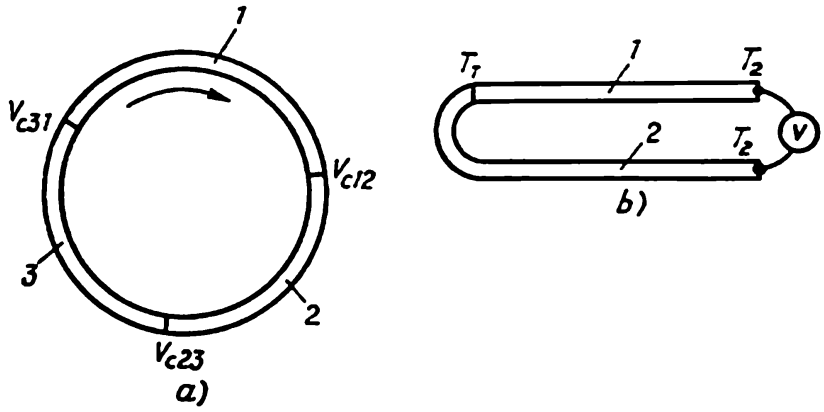

Fig. 10.17. Explication de l'effet de f.é.m. thermoélectrique.

éléments sont portés à une seule et même température. En effet, considérons par exemple un circuit constitué de trois matériaux différents (fig. 10.17, *a*). Pour fixer les idées, nous considérerons comme positives les forces électromotrices qui doivent conduire à l'apparition dans le circuit d'un courant dirigé dans le sens des aiguilles d'une montre. Désignons les travaux d'extraction thermodynamiques des

matériaux considérés respectivement par χ_{01}, χ_{02} et χ_{03}. Alors

$$\begin{aligned} eV_{c12} &= \chi_{01} - \chi_{02}; \\ eV_{c23} &= \chi_{02} - \chi_{03}; \\ eV_{c31} &= \chi_{03} - \chi_{01}. \end{aligned} \tag{10.37}$$

En faisant la somme des premiers et des seconds membres de (10.37), on constate que la somme algébrique de toutes les f.é.m. de contact est nulle. Cette conclusion n'est pas pourtant vraie si les contacts du circuit sont portés à des températures différentes. En effet, faisons varier la température du contact *1-2* par exemple. Comme les énergies de Fermi et, avec elles, les travaux d'extraction thermodynamiques dépendent de la température, la valeur de V_{c12} variera. Quant à toutes les autres valeurs de V_c, elles resteront inchangées. Par conséquent, la somme de V_c cessera d'être nulle et il apparaîtra dans le circuit une différence de potentiel résultante $V_{f.é.m.th}$ qu'on appelle *force thermo-électromotrice.*

La cause d'apparition de force thermo-électromotrice tiest constuée non seulement par la variation des différences de potentiel de contact avec la température mais aussi par l'apparition de flux de diffusion et de thermodiffusion de porteurs de charges libres par suite de la différence de température le long des conducteurs (v. § 6.8). Pour l'étude ultérieure, prenons un exemple bien simple de circuit constitué par deux matériaux et ouvert à l'endroit de l'un des contacts (fig. 10.17, *b*). Pour pouvoir mesurer la f.é.m. thermoélectrique, intercalons dans la coupure de circuit un voltmètre présentant une forte résistance interne. C'est ainsi que nous excluons de l'examen des effets supplémentaires liés au passage du courant dans un circuit présentant des gradients de température.

Comme le montre la théorie, dans un intervalle de température relativement étroit, où la température T_2 d'un des contacts (du contact *2-1* de la figure 10.17, *b*) diffère de la température T_1 de tous les autres contacts, la f.é.m. thermoélectrique

$$V_{f.é.m.th.} = (\alpha_1 - \alpha_2)(T_2 - T_1). \tag{10.38}$$

Les coefficients α_1 et α_2 dépendent des paramètres des matériaux *1* et *2* respectivement. On les appelle *f.é.m. thermoélectriques différentielles.*

Pour un semiconducteur de type n non dégénéré

$$\alpha_n = -\frac{k}{e}\left(r + 2 - \frac{\mu}{kT}\right). \tag{10.39}$$

Ici, r est l'exposant dans la formule exprimant le libre parcours moyen λ en fonction de l'énergie de l'électron E :

$$\lambda \sim E^r. \tag{10.40}$$

L'exposant r dépend du mécanisme de la diffusion : dans la diffusion par des vibrations thermiques acoustiques du réseau $r = 0$, dans la diffusion par des ions d'impureté $r = 2$, etc.

Pour un semiconducteur de type p non dégénéré

$$\alpha_p = \frac{k}{e}\left(r + 2 - \frac{\mu'}{kT}\right), \tag{10.41}$$

où μ' est la distance entre le niveau de Fermi et le sommet de la bande de valence. Pour un métal dont l'énergie dépend paraboliquement du vecteur d'onde

$$\alpha_m = -\frac{\pi^2 k}{3e}(r+1)\frac{kT}{\mu}. \tag{10.42}$$

Evaluons α_m et $\alpha_{n,p}$. Pour des métaux typiques, le cuivre par exemple, $\mu \approx 7$ eV ; à $T = 300$ K $kT = 0{,}025$ eV. En portant cette valeur de kT dans (10.42) et en posant $r = 0$, on obtient $\alpha_m \approx 1$ μV/K, ce qui s'accorde de manière satisfaisante avec l'expérience.

Pour un semiconducteur de type n, par exemple pour le silicium n avec une concentration d'impureté donatrice $N_d = 10^{20}$ m^{-3}, à $T = 300$ K, on a $\mu_n \approx \approx -0{,}3$ eV. En portant cette valeur dans (10.39), on trouve $\alpha_n \approx 1{,}2$ mV/K, ce qui est approximativement de trois ordres de grandeur supérieur à la valeur correspondante pour les métaux. Aussi, pour un circuit comprenant un semiconducteur et un métal, peut-on négliger la part de la f.é.m. thermoélectrique due au métal et considérer que toute la f.é.m. thermoélectrique est due au semiconducteur.

Pour les semiconducteurs à conductibilité mixte, dans lesquels le courant électrique est transporté à la fois par les électrons et les trous, la f.é.m. thermoélectrique est donnée par la relation suivante :

$$\alpha_{n,p} = (\alpha_n u_n n + \alpha_p u_p p)/(n u_n + p u_p). \tag{10.43}$$

Pour des semiconducteurs intrinsèques $n = p = n_i$ et

$$\alpha_i = (\alpha_n u_n + \alpha_p u_p)/(u_n + u_p). \tag{10.44}$$

Dans un semiconducteur intrinsèque $\mu_i \approx -E_g/2$ et, si la largeur de la bande interdite est de l'ordre de 1 eV, les termes $(r + 2)$ figurant dans les expressions (10.39) et (10.41) peuvent être négligés devant $\mu/(kT)$. On obtient alors

$$\alpha_i \approx -\frac{E_g}{2eT}\frac{u_n - u_p}{u_n + u_p}. \tag{10.45}$$

Puisque d'ordinaire $u_n > u_p$, dans un semiconducteur intrinsèque la grandeur α_i a une valeur négative. Dans un semiconducteur à défaut d'électrons le passage à la conduction intrinsèque change le signe positif de la f.é.m. thermoélectrique en signe négatif.

La figure 10.18 montre la variation de la f.é.m. thermoélectrique d'un semiconducteur à défaut d'électrons en fonction de la température. Dans le domaine des basses températures, où l'énergie de Fermi

subit une variation relativement faible avec la température (v. fig. 5.9, *b*), $-\mu'/(kT)$ varie approximativement en raison inverse de la température T. C'est pourquoi la f.é.m. thermoélectrique spécifique $\alpha_p \sim 1/T$ (partie AB de la courbe de la figure 10.18). Dans le domaine de déplétion de l'impureté $-\mu'/(kT) = \ln (N_v/N_a) = \ln N_v - \ln N_a$. Comme $N_v \sim T^{3/2}$, dans ce domaine α_p croît légèrement avec la température (partie BC de la courbe). Enfin, dans le domaine de passage à la conduction intrinsèque μ' devient pratiquement une grandeur constante égale à μ_i' de sorte que α_p diminue de nouveau lorsque la température augmente et la courbe devrait aller comme le montre la partie CD de la figure 10.18. Pourtant du fait que dans ce domaine la concentration des électrons augmente et devient pratiquement égale à celle des trous, la f.é.m. thermoélectrique tombe brusquement, passe par zéro et devient négative (partie CE de la courbe).

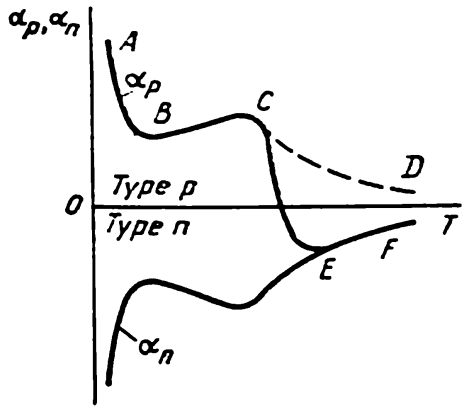

Fig. 10.18. Variation avec la température de la f.é.m. thermoélectrique des semiconducteurs de type *p* et de type *n*.

Lorsque la température augmente encore, le module de la f.é.m. thermoélectrique diminue suivant la formule (10.45) (la courbe EF pour un semiconducteur intrinsèque).

La partie inférieure de la figure 10.18 montre la courbe de variation de la f.é.m. thermoélectrique α_n d'un semiconducteur de type n en fonction de la température. Dans le domaine de la conductibilité intrinsèque cette courbe se confond avec la courbe pour un semiconducteur de type p.

L'effet thermoélectrique a trouvé de nombreuses applications pratiques y compris en radio-électronique. Il permet de convertir l'énergie thermique directement en énergie électrique, ce qui est ntilisé dans les *générateurs thermoélectriques*. La théorie de tels générateurs a été développée par le physicien soviétique A. Ioffé. Suivant cette théorie le rendement de conversion de l'énergie thermique en énergie électrique se détermine par la quantité $\alpha^2\sigma/\lambda$, où λ est le coefficient de conductibilité thermique du semiconducteur ; σ la conductivité électrique.

On peut expliciter cela qualitativement en partant des considérations suivantes. Dans les générateurs thermoélectriques on cherche à obtenir une différence de température aussi grande que possible entre les extrémités chaude et froide du semiconducteur pour une dépense aussi petite que possible de l'énergie thermique. A cet effet, il est avantageux d'utiliser des semiconducteurs possédant une fai-

ble conductibilité thermique parce que la diminution de la transmission de chaleur de l'extrémité chaude à l'extrémité froide grâce à l'allongement du semiconducteur a pour résultat une augmentation de la résistance interne du générateur thermoélectrique et une chute du rendement. Pour cette même raison, il est avantageux que la conductivité électrique du semiconducteur soit maximale. Vu que lorsque le degré de dopage du semiconducteur augmente, α décroît, alors que λ et σ croissent, il existe pour chaque semiconducteur un degré de dopage optimal qui assure la valeur maximale du rapport $\alpha^2\sigma/\lambda$ et donc le rendement maximal.

Les premiers générateurs thermoélectriques ont été élaborés avant la Grande Guerre Nationale et utilisé pendant la guerre pour l'alimentation des appareils radio-électriques. En 1953 a été construit le premier générateur thermoélectrique industriel ayant une puissance de 1 kW. De nos jours on élabore des générateurs pour une puissance de plusieurs centaines de kilowatts.

Au milieu des années 70, on a vu apparaître des générateurs utilisant la chaleur dégagée lors de la désintégration radioactive des éléments chimiques. Comme exemple d'un tel générateur, on peut indiquer l'installation « Bêta-1 » ayant une puissance de 150 à 200 W utilisant l'isotope de cérium-144 radioactif. Elle est destinée à assurer l'alimentation des appareils radio-électroniques des stations radiométéorologiques automatiques, des satellites artificiels de la Terre, etc.

En 1964, on a construit un réacteur-convertisseur atomique expérimental « Romachka » de 500 W de puissance qui transforme l'énergie thermique directement en énergie électrique.

Le rendement des générateurs thermoélectriques à semiconducteurs atteint actuellement 15% et, à ce qu'il paraît, il dépassera 20 % dans le plus proche avenir.

L'effet thermoélectrique est également utilisé pour la détermination des températures (couples thermoélectriques) et pour d'autres mesures qui peuvent être ramenées à la mesure de la température. Dans les photorécepteurs thermiques (thermoéléments) la lumière qui est absorbée échauffe une surface noircie de réception à laquelle est reliée la soudure de couple thermoélectrique. D'après la valeur de la f.é.m. thermoélectrique qui prend naissance, on peut déterminer la puissance du flux lumineux. Dans les *vacuomètres*, on fait passer un courant électrique fixé dans un conducteur métallique au milieu duquel est reliée la soudure d'un couple thermoélectrique. La température de la soudure dépend de la conductibilité thermique du gaz environnant. Or, cette dernière se détermine par la pression du gaz. Aussi, en mesurant la valeur de la f.é.m. thermoélectrique qui apparaît, peut-on déterminer la pression du gaz. Cette méthode permet de mesurer commodément les pressions dans la gamme de 10^{-1} à 10 Pa.

Dans la technique des mesures de paramètres électrophysiques des semiconducteurs, l'effet thermoélectrique est utilisé pour la détermination du type de conductibilité prédominante (d'après le signe de la f.é.m. thermoélectrique) et de la largeur de la bande interdite [par la formule (10.45)].

§ 10.7. Effet Peltier

L'*effet Peltier* consiste en ce que lorsqu'on fait passer un courant électrique à travers un contact de deux matériaux différents, il se produit, en plus du dégagement de la chaleur Joule, encore le dégagement ou l'absorption de la chaleur dont la quantité Q_P est proportionnelle à la charge It ayant traversé le contact :

$$Q_P = \pm \mathrm{P}It. \tag{10.46}$$

Le coefficient de proportionnalité P est appelé *coefficient de Peltier.*

Expliquons la nature de cet effet sur l'exemple d'un contact neutre entre un semiconducteur et un métal. On appelle *contact neutre* le contact de deux matériaux possédant le même travail d'extraction.

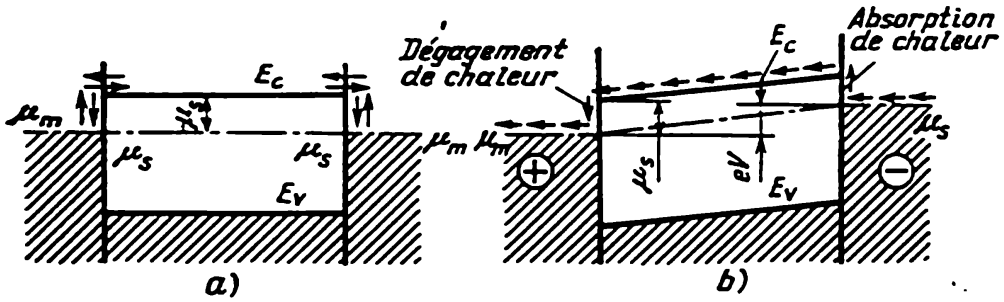

Fig. 10.19. Explication de l'effet Peltier :
a — diagramme énergétique d'un circuit métal-semiconducteur-métal à contacts neutres en équilibre ; *b* — idem lors du passage d'un courant.

Dans un tel contact il n'existe ni les couches d'enrichissement, ni les couches de déplétion, et les bandes d'énergie ne sont pas incurvées. La figure 10.19, *a* montre le diagramme des énergies d'une structure métal-semiconducteur-métal avec un contact neutre entre un métal et un semiconducteur de type *n*. A l'état d'équilibre, les niveaux de Fermi du métal μ_m et du semiconducteur μ_s se situent à une même hauteur et le niveau inférieur de la bande de conduction du semiconducteur se situe plus haut que le niveau de Fermi du métal de $-\mu_s$ de sorte que les électrons passant du métal dans le semiconducteur ont à franchir une barrière de potentiel haute de $-\mu_s$.

Appliquons à une telle structure une différence de potentiel extérieure, comme il est montré sur la figure 10.19, *b*. Sous l'action de cette différence de potentiel, dans le semiconducteur où chute pratiquement toute la tension appliquée, il se produit, dans le sens

du contact de gauche vers le contact de droite, une montée progressive d'une hauteur eV du niveau de Fermi, du niveau inférieur de la bande de conduction et des autres niveaux d'énergie. Dans le système il naît un courant dirigé de gauche à droite ; le flux d'électrons (montré sur la figure 10.19, *b* par des flèches) est dirigé de droite à gauche.

L'examen de la figure 10.19, *b* montre que les électrons qui passent dans le contact de gauche du semiconducteur dans le métal possèdent dans le semiconducteur une énergie de $-\mu_n + \langle E_n \rangle$ plus grande que dans le métal, où $-\mu_n$ est la hauteur de la barrière que franchissent les électrons passant du semiconducteur dans le métal ; $\langle E_n \rangle$ l'énergie moyenne des électrons dans le semiconducteur participant à la création du courant. Elle n'est pas égale à l'énergie thermique moyenne des électrons $^3/_2 kT$, comme il semblerait à première vue, parce que le rôle relatif des électrons rapides dans la formation du courant électrique est plus grand que celui des électrons lents. Pour un gaz électronique non dégénéré le calcul montre que

$$\langle E_n \rangle = (r + 2)\, kT, \tag{10.47}$$

où r est l'exposant dans la relation (10.40).

Ainsi, chaque électron passant dans le contact de gauche du semiconducteur dans le métal transporte une énergie excédentaire

$$\Delta E = -\mu_n + (r + 2)\, kT, \tag{10.48}$$

qui se dégage dans ce contact sous forme de chaleur supplémentaire (chaleur de Joule). C'est cette chaleur qui représente la chaleur de Peltier. Si l'on divise (10.48) par la charge de l'électron e, on obtient le coefficient de Peltier

$$|\pi| = \frac{k}{e}\left(r + 2 + \frac{\mu_n}{kT}\right) T. \tag{10.49}$$

Dans le contact de droite, les électrons passent du métal dans le semiconducteur en franchissant la barrière de potentiel haute de $-\mu_n$. En outre, pour établir un équilibre entre ces électrons et le gaz électronique existant dans le semiconducteur, les électrons doivent acquérir encore une énergie $\langle E_n \rangle$; ils empruntent toute cette énergie au réseau cristallin en le refroidissant au contact de droite.

Le phénomène de refroidissement du contact lorsqu'il est traversé par un courant électrique est particulièrement important parce qu'il permet de réaliser des *installations frigorifiques* : les réfrigérateurs ménagers, les dispositifs pour le refroidissement de l'appareillage radioélectronique de bord, les microréfrigérateurs pour usages biologiques, les thermostats thermoélectriques variés, les platines du microscope avec réchauffement et refroidissement, etc. Ce qui est bien commode c'est la possibilité de transformer un réfrigérateur en réchauffeur et, inversement, un réchauffeur en réfrigérateur en changeant le sens du courant.

Il est facile de voir que l'effet Peltier est l'inverse de l'effet de f.é.m. thermoélectrique (*effet Seebeck*). Dans le premier cas, le passage d'un courant dans un circuit conduit à l'apparition aux contacts d'une différence de température et dans le second cas la création d'une différence de température aux contacts a pour effet l'apparition dans le circuit d'une f.é.m. thermoélectrique et donc d'un courant électrique. La comparaison de (10.47) et (10.39) ainsi que l'étude thermodynamique de ces phénomènes montre qu'entre le coefficient de Peltier et la f.é.m. thermoélectrique spécifique il existe une relation simple suivante :

$$P = \alpha T. \tag{10.50}$$

Les matériaux les plus efficaces pour la réalisation aussi bien des réfrigérateurs thermoélectriques que des générateurs thermoélectriques sont ceux dont la valeur de $\alpha^2\sigma/\lambda$ est maximale. Pour le refroidissement thermoélectrique il faut avoir un matériau présentant des valeurs élevées du coefficient de Peltier et de la conductivité électrique. Cette dernière exigence tient à ce qu'en plus de la chaleur de Peltier il y a dégagement de la chaleur de Joule et pour que l'effet de réchauffement par la chaleur de Joule ne dépasse pas l'effet de refroidissement, il faut utiliser un matériau à bonne conductibilité électrique. D'un autre côté, pour une même quantité de chaleur dégagée par l'effet Peltier à l'un des contacts et absorbée à l'autre contact, la différence de température entre les contacts sera d'autant plus grande que la transmission de chaleur de l'extrémité chaude du semiconducteur à l'extrémité froide sera plus faible, c'est-à-dire que le coefficient de conductibilité thermique est moins élevé.

CHAPITRE 11

DIODES À SEMICONDUCTEURS

Une des classes les plus larges d'appareils électroniques à solides est constituée par des diodes à semiconducteurs destinées à effectuer les transformations non linéaires des signaux électriques (redressement, détection, modulation, génération, etc.), la conversion de l'énergie électrique en rayonnement (diodes électroluminescentes, lasers) et inversement, la conversion du rayonnement en énergie électrique (convertisseurs photo-électriques, piles solaires) et beaucoup d'autres fonctions. La majeure partie des diodes utilise les propriétés uniques du contact entre deux semiconducteurs dopés de différents types. Un tel contact est appelé *jonction électron-trou* ou, plus couramment, jonction *p-n*. Le présent chapitre est consacré à l'étude des propriétés de la jonction *p-n* et des principes physiques de fonctionnement de principaux types de diodes utilisant ces propriétés.

§ 11.1. Etat d'équilibre d'une jonction *p-n*

Préparation d'une jonction *p-n*. Il est pratiquement impossible d'obtenir une jonction *p-n* par une mise en contact directe de deux semiconducteurs parce que leurs surfaces, même nettoyées avec le plus grand soin, contiennent une quantité énorme d'impuretés, d'encrassements et de défauts de toutes natures qui modifient radicalement les propriétés des semiconducteurs. Aussi un progrès remarquable dans la mise en valeur de la jonction *p-n* n'a-t-il été réalisé que lorsqu'on a appris à l'obtenir sous la forme d'une frontière intérieure dans un semiconducteur monocristallin.

Considérons les méthodes les plus employées de nos jours pour la préparation d'une jonction *p-n*.

Méthode de la diffusion. Une jonction *p-n* peut être obtenue par la diffusion d'une impureté acceptrice dans un semiconducteur de type *n* ou d'une impureté donatrice dans un semiconducteur de type *p*. La diffusion de l'impureté s'effectue à des températures élevées (1000 à 1300 °C) soit à partir d'un diffusant solide déposé sur la surface d'une plaquette semiconductrice, soit à partir d'un gaz-porteur contenant l'impureté de dopage requise, qu'on fait circuler sous forme de flux au-dessus des plaquettes. La figure 11.1, *a* représente schématiquement la répartition d'une impureté acceptrice après sa diffusion dans un semiconducteur de type *n* (N_d est la concentration des donneurs dans la plaquette semiconductrice). Près de la surface de la plaquette le type de conduction est déterminé par l'impureté acceptrice dont la concentration est ici $N_a > N_d$. A mesure qu'on s'éloi-

gne de la surface de la plaquette, la concentration N_a diminue et, à une certaine profondeur x_0, elle devient égale à la concentration de l'impureté donatrice dans la plaquette. C'est dans ce plan que se forme une jonction *p-n*. En aval de cette jonction $N_a < N_d$ et diminue rapidement à mesure qu'on s'éloigne de la jonction *p-n*. De telles jonctions sont dites *graduelles*.

Méthode de l'épitaxie. Cette technique consiste à déposer sur un support monocristallin, par exemple de silicium de type *n*, une couche monocristalline de type *p*. Une jonction *p-n* se forme à la frontière entre le support et cette couche. Le dépôt lui-même de la couche épitaxiale peut se faire à partir d'une phase gazeuse ou liquide.

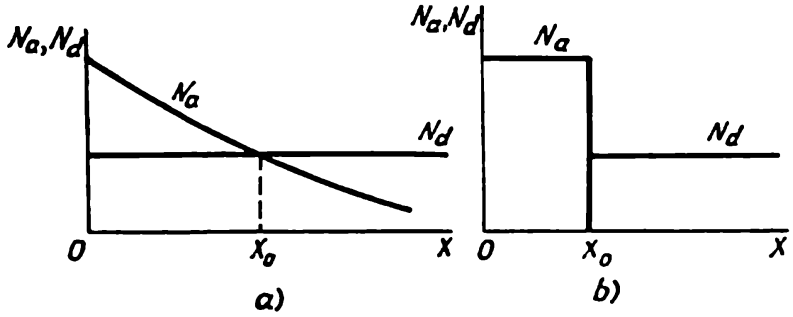

Fig. 11.1. Répartition des impuretés dans des jonctions *p-n* diffusée (*a*) et épitaxiale (*b*).

Comme exemple d'épitaxie en phase gazeuse, on peut indiquer le dépôt d'un film de silicium sur des supports en silicium à partir du tétrachlorure de silicium $SiCl_4$. A cet effet, dans un tube de quartz contenant des plaquettes de silicium échauffées à une température d'environ 1200 °C, on fait passer un courant d'hydrogène saturé en vapeurs de $SiCl_4$. En passant au-dessus des plaquettes de silicium, le $SiCl_4$ se réduit jusqu'au silicium atomique ($SiCl_4 + 2H_2 \rightarrow 4HCl + Si$) qui se dépose sur les plaquettes sous la forme d'un film épitaxial monocristallin de silicium. Pour le dopage des films avec une impureté acceptrice, on ajoute au $SiCl_4$ du bromure de bore BBr_3 et pour le dopage avec une impureté donatrice, on ajoute du chlorure de phosphore PCl_3.

Comme exemple d'épitaxie en phase liquide, on peut indiquer le dépôt de couches épitaxiales de GaAs sur un substrat de GaAs à partir d'une solution d'arséniure de gallium dans le gallium métallique fondu. La quantité de GaAs dans la solution est réglée de telle sorte qu'à une température supérieure à T_1 la solution soit non saturée et à une température inférieure à T_1 sursaturée. On échauffe la solution jusqu'à la température T_1 et on la dépose sur des plaquettes d'arséniure de gallium portées à la même température. Après cela, la température à l'intérieur du four est abaissée au-dessous de T_1 de sorte que des films monocristallins de GaAs commencent à croître

sur la surface du support à partir de la solution d'arséniure de gallium. Pour le dopage de tels films on introduit dans la solution des additifs de dopage convenables.

La figure 11.1, *b* montre une répartition idéalisée d'une impureté acceptrice (de concentration N_a) dans un film épitaxial déposé sur un support dopé avec une impureté donatrice (de concentration N_d). Une telle répartition correspond à une *jonction p-n* dite *abrupte*. En réalité, les fronts de la répartition des impuretés dopantes sont un peu flous par suite du processus de diffusion. Plus basse est la température et plus court est le temps au cours duquel s'effectue l'épitaxie, plus abruptes sont les jonctions *p-n* obtenues.

Ces derniers temps, on développe la technique d'épitaxie dans un vide poussé (10^{-7} à 10^{-8} Pa) quand le support est bombardé par un flux d'atomes de substance à déposer. Si l'on dépose un composé (par exemple GaAs ou $Ga_xAl_{1-x}As$), chacun des constituants est envoyé sur le support depuis sa source distincte. Cette méthode, qui a reçu le nom d'*épitaxie par faisceau moléculaire*, permet d'obtenir des films de composition contrôlable, de très petite épaisseur allant jusqu'à quelques nanomètres.

Enfin, on emploie de plus en plus couramment la méthode de préparation de jonctions *p-n* à l'aide du *dopage par implantation d'ions*, qui consiste à envoyer sur un support semiconducteur de type donné de conduction un faisceau d'ions d'impureté de type de conduction contraire. L'énergie des ions dans le faisceau se chiffre par des dizaines et des centaines de milliers d'électrons-volts. Ces ions restent dans la couche superficielle du semiconducteur, mais à l'état électriquement inactif il est vrai (en position intersticielle). Pour rendre l'impureté implantée électriquement active (c'est-à-dire pour la faire passer dans les nœuds du réseau cristallin), il faut procéder au recuit, c'est-à-dire à l'échauffement des plaquettes pendant un certain temps (habituellement durant quelques dizaines de minutes). Le recuit permet non seulement d'activer l'impureté mais aussi de rétablir le réseau cristallin perturbé lors de l'implantation. Si le recuit est effectué à une température pas trop élevée (pour le silicium, inférieure à 1000 °C), la jonction *p-n* obtenue sera suffisamment abrupte.

Jonction *p-n* à l'état d'équilibre. Supposons que la frontière intérieure entre deux régions du semiconducteur présentant des types de conduction contraires est constituée par le plan *MM* (fig. 11.2, *a*) : à gauche de ce plan se trouve un semiconducteur de type *p* (par exemple le germanium *p*) avec la concentration des accepteurs N_a et à droite un semiconducteur de type *n* (le germanium *n*) avec la concentration des donneurs N_d. Admettons pour simplifier que $N_a = N_d = 10^{22}$ m^{-3}. La figure 11.2, *b* montre la variation de la concentration des atomes accepteurs et donneurs lors du déplacement suivant l'axe des x perpendiculaire au plan *MM*. Au point *O* appartenant à ce plan, la concentration des accepteurs N_a s'annule soudai-

nement tandis que la concentration des donneurs varie brutalement de zéro à N_d.

Pour la région n les porteurs majoritaires sont les électrons et pour la région p, les trous. Les porteurs majoritaires sont produits presque entièrement par ionisation des impuretés donatrices et acceptrices.

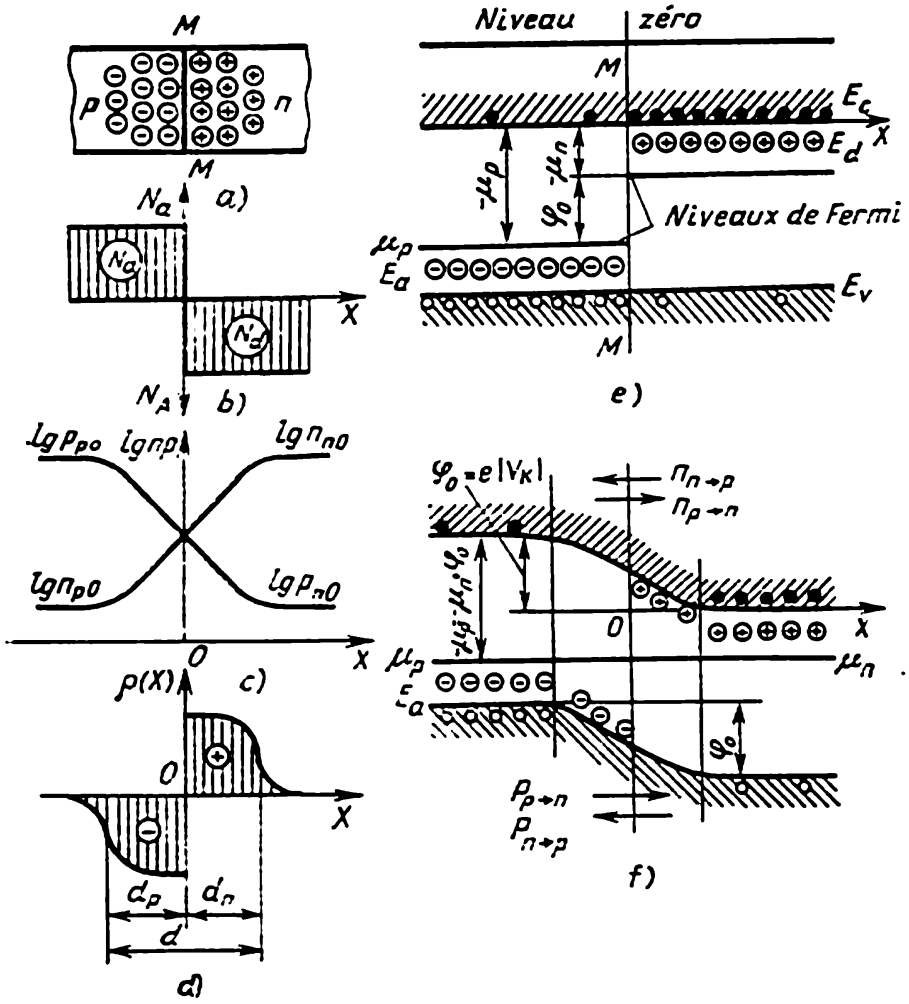

Fig. 11.2. Jonction p-n en équilibre.

Si les températures ne sont pas trop basses, ces impuretés sont ionisées pratiquement en totalité de sorte qu on peut considérer que la concentration n_{n0} des électrons dans la région n est pratiquement égale à la concentration des atomes donneurs N_d (c'est-à-dire que $n_{n0} \approx \approx N_d$) et la concentration p_{p0} des trous dans la région p est égale à la concentration des atomes accepteurs N_a (c'est-à-dire que $p_{p0} \approx \approx N_a$).

En plus des porteurs majoritaires ces régions contiennent des porteurs minoritaires : la région n les trous (p_{n0}) et la région p les électrons (n_{p0}). Leurs concentrations peuvent être calculées à l'aide

de la loi d'action de masse (5.36):

$$n_{n0}p_{n0} = p_{p0}n_{p0} = n_i^2, \tag{11.1}$$

où n_i est la concentration des porteurs dans un semiconducteur intrinsèque. Pour $n_{n0} = p_{p0} = 10^{22}$ m^{-3} et $n_i = 10^{19}$ m^{-3} (dans le Ge) on obtient $p_{n0} = n_{p0} = 10^{16}$ m^{-3}.

On voit que la concentration des trous dans la région p est supérieure de six ordres de grandeurs à leur concentration dans la région n, de même la concentration des électrons dans la région n est de six ordres de grandeur supérieure à leur concentration dans la région p. Une telle différence entre les concentrations des porteurs de même type dans les régions en contact du semiconducteur fait naître des *flux de diffusion* des électrons de la région n dans la région p et des trous de la région p dans la région n. Les électrons, ayant passé de la région n dans la région p, se recombinent au voisinage de la surface de séparation de ces régions avec les trous de la région p, de même que les trous, ayant passé de la région p dans la région n, s'y recombinent avec les électrons de cette région. Il en résulte que dans la couche adjacente au contact de la région n il ne reste pratiquement pas d'électrons libres et il s'y forme une *charge d'espace positive immobile* de donneurs ionisés (fig. 11.2, *a*).

Dans la couche de contact de la région p il ne reste pratiquement pas de trous et il s'y forme une charge d'espace négative, immobile, d'accepteurs ionisés. La figure 11.2, *c* montre la répartition des porteurs de charges libres dans la région de la jonction p-n et la figure 11.2, *d* représente des charges d'espace immobiles produites dans cette jonction (ρ étant la masse volumique de ces charges). Désignons par d_p et d_n les épaisseurs des couches de la charge d'espace respectivement dans la région p et dans la région n et par d l'épaisseur totale de la couche de charge d'espace.

Les charges d'espace immobiles engendrent dans la jonction p-n un *champ électrique de contact*, de différence de potentiel V_c, qui est localisé dans la région de la jonction et pratiquement ne s'exerce pas hors de ses limites. C'est pourquoi, en dehors de cette couche, où le champ électrique est nul, les porteurs de charges libres sont animés, comme précédemment, d'un mouvement désordonné et le nombre de porteurs qui viennent tomber chaque seconde sur la couche de charge d'espace dépend uniquement de leur concentration et de la vitesse de l'agitation thermique.

Si un porteur minoritaire (un électron de la région p ou un trou de la région n) pénètre dans la couche de charges d'espace, le champ électrique de contact l'« attrape » et le fait passer à travers cette couche. Dans ce cas, la probabilité de recombinaison du porteur pendant le temps de transit de la région de charge d'espace peut être souvent (dans des jonctions p-n dites *étroites*) négligée. Alors, chaque porteur minoritaire tombant sur la jonction p-n la traverse. Par contre, les

porteurs de courant majoritaires (les électrons de la région n et les trous de la région p) ne peuvent traverser la couche de charges d'espace que si l'énergie cinétique de leur mouvement le long de l'axe des x est suffisante pour surmonter la différence de potentiel de contact, c'est-à-dire si cette énergie est plus grande que $e \mid V_c \mid$. Aussi, dès que les charges d'espace se forment près de la surface frontière MM, les flux de porteurs majoritaires traversant cette frontière diminuent-ils. Pourtant si ces flux dépassent encore les flux opposés de porteurs minoritaires qui restent inchangés, la charge d'espace ira croissant. Cette croissance continue tant que les flux décroissants de porteurs majoritaires ne s'égalisent avec les flux de porteurs minoritaires. C'est ainsi que s'établit un *état d'équilibre dynamique* de la jonction p-n.

Ainsi, à l'état d'équilibre, la jonction p-n est traversée par des flux égaux des électrons venant de la région n dans la région p et de la région p dans la région n, ainsi que par des flux égaux des trous venant de la région p dans la région n et de la région n dans la région p. L'égalité de ces flux est assurée par l'existence d'une différence de potentiel de contact déterminée V_c qui réduit les flux de porteurs majoritaires traversant la jonction jusqu'au niveau des flux de porteurs minoritaires. En se servant de la distribution de Maxwell des particules suivant les vitesses, on peut montrer que le nombre de porteurs de courant libres venant frapper par unité de temps la couche de charge d'espace est égal à 1/4 du produit de la concentration des porteurs par la vitesse moyenne $\langle v \rangle$ de leur agitation thermique et par la surface S de la jonction p-n. Par exemple, le nombre des électrons frappant la jonction par unité de temps du côté de la région p est égal à $^1/_4 n_{p0} \langle v_n \rangle S$ et le nombre des électrons attaquant la jonction du côté de la région n vaut $^1/_4 n_{n0} \langle v_n \rangle S$. Le champ de contact diminue les flux des porteurs majoritaires surmontant la différence de potentiel V_c de $\exp [e \mid V_c \mid /(kT)]$ fois de sorte que le flux des électrons venant de la région n dans la région p du semiconducteur devient égal à

$$^1/_4\, n_{n0} \langle v_n \rangle S e^{-e \mid V_c \mid /(kT)} = {}^1/_4\, n_{p0} \langle v_n \rangle S. \tag{11.2}$$

On en tire

$$n_{n0} e^{-e \mid V_c \mid /(kT)} = n_{p0}; \tag{11.3}$$

$$e \mid V_c \mid = kT \ln \frac{n_{n0}}{n_{p0}} = kT \ln (n_{n0} p_{p0} / n_i^2). \tag{11.4}$$

La dernière expression a été obtenue à l'aide de la loi d'action de masse (11.1). En égalant entre eux les flux opposés des trous traversant la jonction p-n, on obtient

$$p_{p0} e^{-e \mid V_c \mid /(kT)} = p_{n0}; \tag{11.5}$$

$$e \mid V_c \mid = kT \ln \frac{p_{p0}}{p_{n0}} = kT \ln (n_{n0} / p_{p0} / n_i^2). \tag{11.6}$$

Le rapprochement entre les expressions (11.4) et (11.6) montre que les flux opposés des électrons et des trous s'égalisent pour une même différence de potentiel de contact V_c. De (11.4) et (11.6) il résulte aussi que la barrière de potentiel est d'autant plus haute que la différence entre les concentrations des porteurs de même signe dans les régions n et p est plus grande. Lorsqu'on met en contact deux semiconducteurs intrinsèques ($n_{n0} = p_{p0} = n_i$), la barrière de potentiel ne se forme pas. C'est pourquoi la jonction p-n disparaît lorsque la température de conductibilité intrinsèque est atteinte.

La figure 11.2, *e* montre le diagramme de bandes d'énergie des régions p et n d'un semiconducteur à l'instant où ces régions sont mis, par la pensée, en contact, c'est-à-dire avant l'établissement de l'équilibre entre ces régions. Les niveaux énergétiques sont représentés par des droites horizontales. Cela exprime le fait que l'énergie d'un électron occupant un niveau donné, par exemple le niveau inférieur de la bande de conduction, est la même en tous les points du semiconducteur. Après l'établissement de l'équilibre il se forme une jonction p-n avec une barrière de potentiel pour les porteurs majoritaires $\varphi_0 = e\,|\,V_c\,|$. Les électrons passant de la région n dans la région p, franchissant cette barrière, augmentent leur énergie potentielle de $\varphi_0 = e\,|\,V_c\,|$. C'est pourquoi tous les niveaux énergétiques d'un semiconducteur s'incurvent dans la région de la jonction p-n pour remonter de φ_0 comme l'indique la figure 11.2, *f*. Dans ce cas les niveaux de Fermi μ_n et μ_p s'établissent à une même hauteur. En effet, pour un semiconducteur non dégénéré

$$n_{n0} = N_c e^{\mu_n/(kT)}; \quad n_{p0} = N_c e^{\mu_p/(kT)}, \tag{11.7}$$

où N_c est la densité effective d'états près du niveau inférieur de la bande de conduction ; μ_n et μ_p sont les niveaux de Fermi dans les régions n et p respectivement. En portant (11.7) dans (11.4), on obtient

$$\varphi_0 = \mu_n - \mu_p. \tag{11.8}$$

Pour déterminer la forme de la fonction $\varphi(x)$ qui caractérise la variation de l'énergie potentielle d'un électron lorsqu'il passe de la région n dans la région p (ou d'un trou passant de la région p dans la région n) on peut se servir de l'*équation de Poisson*

$$\frac{d\varphi^2}{dx^2} = \frac{e}{\varepsilon\varepsilon_0}\rho(x), \tag{11.9}$$

dans laquelle $\rho(x)$ représente la densité volumique des charges créatrices du champ. Nous supposerons que les niveaux donneurs et accepteurs sont complètement ionisés et que pratiquement tous les électrons ont quitté la couche d_n et tous les trous ont quitté la couche d_p. Alors, dans la région n, pour $0 < x < d_n$, la densité volumique des charges $\rho(x) \approx eN_d \approx en_{n0}$ et dans la région p, pour $-d_p < x < 0$,

elle est égale à $\rho(x) \approx -eN_a = -ep_{p0}$. En portant ces valeurs dans (11.9), on obtient

$$\frac{d^2\varphi}{dx^2} = \frac{e^2}{\varepsilon\varepsilon_0} N_d \text{ (pour } x > 0), \tag{11.10}$$

$$\frac{d^2\varphi}{dx^2} = -\frac{e^2}{\varepsilon\varepsilon_0} N_a \text{ (pour } x < 0). \tag{11.11}$$

Vu qu'aux distances $x \geqslant d_n$ et $x \leqslant -d_p$ le champ électrique de contact dans le semiconducteur est nul, les conditions aux limites imposées à ces équations sont les suivantes:

$$\varphi(x)|_{x=d_n} = 0, \quad \varphi(x)|_{x=-d_p} = \varphi_0; \tag{11.12}$$

$$\left.\frac{d\varphi}{dx}\right|_{x=d_n} = 0, \quad \left.\frac{d\varphi}{dx}\right|_{x=-d_p} = 0. \tag{11.13}$$

La résolution des équations (11.10) et (11.11) avec les conditions aux limites (11.12) et (11.13) conduit aux résultats suivants:

$$\varphi = \frac{e^2}{2\varepsilon\varepsilon_0} N_d (d_n - x)^2 \text{ (pour } 0 < x < d_n); \tag{11.14}$$

$$\varphi = \varphi_0 - \frac{e^2}{2\varepsilon\varepsilon_0} N_a (d_p + x)^2 \text{ (pour } -d_p < x < 0); \tag{11.15}$$

$$d_0 = d_n + d_p = \left(\frac{2\varepsilon\varepsilon_0\varphi_0}{e^2} \frac{N_d + N_a}{N_d N_a}\right)^{1/2} = \left(\frac{2\varepsilon\varepsilon_0\varphi_0}{e^2} \frac{n_{n0} + p_{p0}}{n_{n0} p_{p0}}\right)^{1/2}; \tag{11.16}$$

$$d_n/d_p = N_a/N_d. \tag{11.17}$$

Les équations (11.14) et (11.15) montrent que la hauteur de la barrière de potentiel $\varphi(x)$ est une fonction quadratique de la coordonnée x. Conformément à (11.16), la couche de charge d'espace est d'autant plus épaisse que la concentration des porteurs majoritaires égale à la concentration de l'impureté dopante est plus faible. La profondeur de pénétration du champ de contact est plus grande dans la région du semiconducteur qui est dopée plus faiblement. Pour $N_d \ll N_a$ la couche se localise pratiquement en totalité dans la région n:

$$d_0 \approx d_n = \left(\frac{2\varepsilon\varepsilon_0\varphi_0}{e^2 N_d}\right)^{1/2} = \left(\frac{2\varepsilon\varepsilon_0\varphi_0}{e^2 n_{n0}}\right)^{1/2}. \tag{11.18}$$

C'est ainsi que pour le germanium à $N_d = 10^{21}$ m^{-3}, $N_a = 10^{23}$ m^{-3}, on obtient $d_p = 8 \cdot 10^{-9}$ m et $d_n = 8 \cdot 10^{-7}$ m.

Le calcul de l'épaisseur de la couche de charge d'espace est effectué pour une jonction p-n abrupte dans laquelle la concentration des impuretés varie par saut (voir fig. 11.1, *b* et 11.2). Dans le cas des jonctions diffusées, la concentration de l'impureté dans la jonction varie graduellement (v. fig. 11.1, *a* et 11.3). En admettant que cette

variation suit une loi linéaire, on obtient, à partir de la solution de l'équation de Poisson, l'expression suivante pour l'épaisseur de la couche de charge d'espace:

$$d_0 = [12\varepsilon\varepsilon_0\varphi_0/(e^2a)]^{1/3}, \quad (11.19)$$

où a est un coefficient de proportionnalité figurant dans l'expression de la loi de variation linéaire de la concentration de l'impureté non compensée $N_a - N_d$ dans la jonction: $N_a - N_d = a(x - x_0)$ (v. fig. 11.1,*a*).

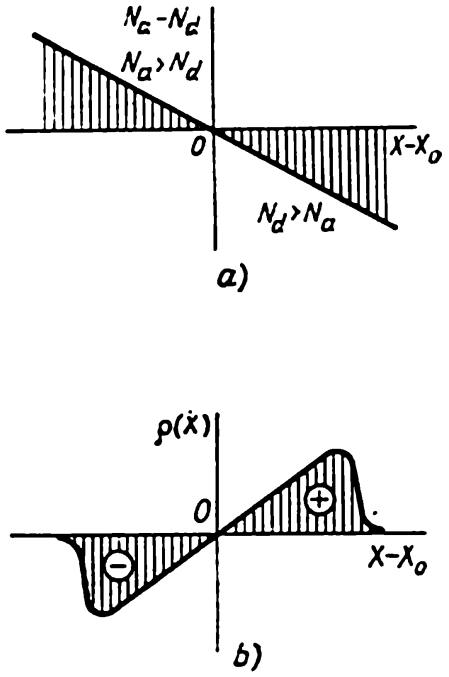

Fig. 11.3. Jonction *p-n* graduelle:
a — répartition d'une impureté non compensée dans la région de jonction; *b* — répartition de densité de la charge d'espace.

§ 11.2. Capacité de barrière d'une jonction *p-n*

La jonction *p-n* est l'élément essentiel des appareils à semiconducteurs les plus variés dont le plus simple (d'après sa constitution) est la *diode*. Dans la diode on relie aux régions *n* et *p* du semiconducteur, à l'aide de contacts ohmiques, des électrodes métalliques qui permettent le passage d'un courant à travers l'appareil. La résistance d'un contact ohmique de bonne qualité est linéaire et de petite valeur par rapport à la résistance de la jonction *p-n* de la diode.

Examinons les phénomènes qui interviennent dans une diode soumise à une différence de potentiel appliquée depuis une source de polarisation extérieure. Si le pôle positif de la source est relié à la région *n* et le pôle négatif à la région *p* la polarisation est dite *inverse* (fig. 11.4).

Après la fermeture de la clé K, des charges affluent par le circuit extérieur vers les contacts ohmiques de la diode et engendrent en profondeur de la diode un champ électrique $\mathscr{E}_{\text{int}}$. Ce champ provoque une dérive des porteurs majoritaires dans des sens indiqués par des flèches sur la figure 11.4, *a*. Ainsi, toute la masse des électrons de la région *n* et des trous de la région *p* s'éloigne de la jonction *p-n*, en mettant à nu de nouvelles couches de donneurs et d'accepteurs ionisés, c'est-à-dire en élargissant la région de charge d'espace jusqu'à des dimensions $d_0 + \Delta d$ (fig. 11.4, *b*).

Le passage des électrons et des trous vers les contacts ohmiques se poursuit jusqu'à l'instant où ils compensent complètement les charges produites par la source de f.é.m. extérieure. Après cela tou-

te la tension appliquée V chute dans la jonction p-n dont la résistance est supérieure de beaucoup d'ordres de grandeur à celle des régions

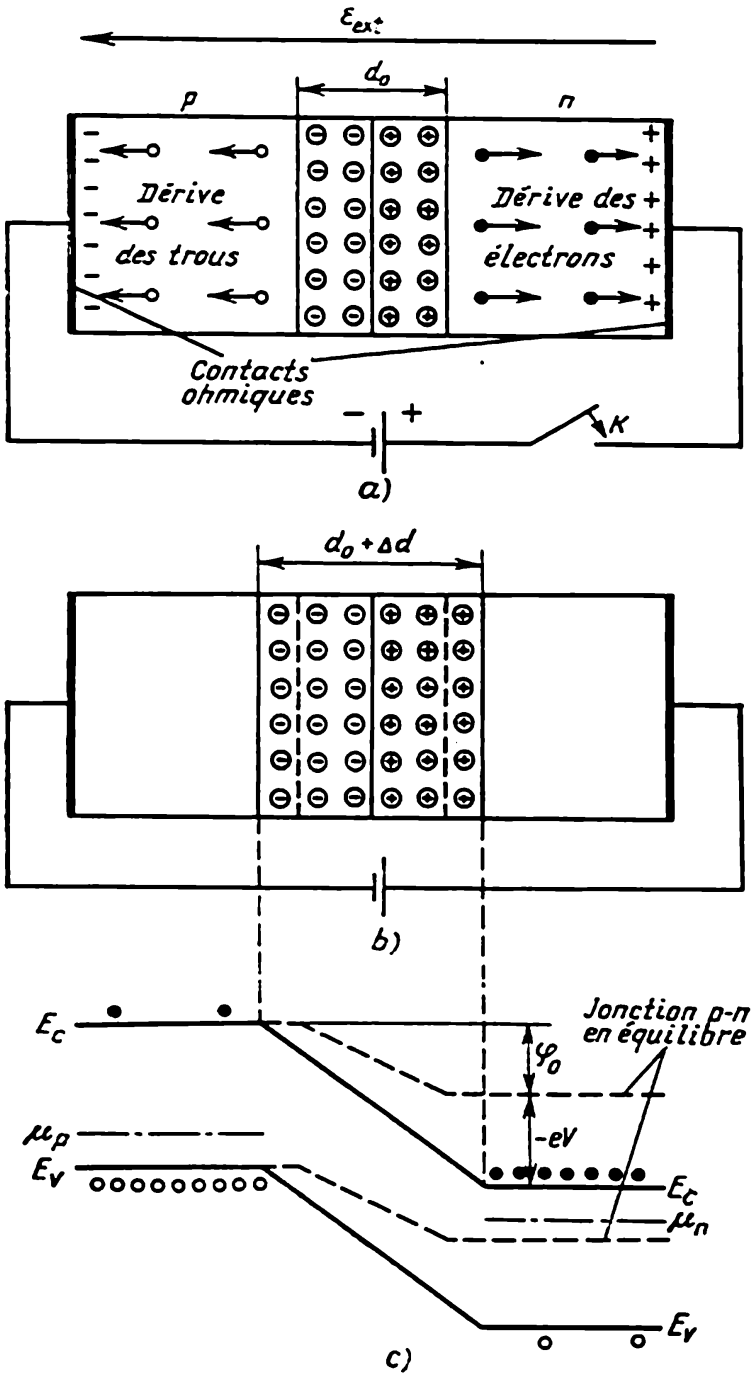

Fig. 11.4. Processus transitoires (*a*) et état stationnaire (*b*, *c*) d'une diode polarisée en sens inverse.

n et p. On convient de considérer la tension V en polarisation inverse comme négative et d'appeler la courbe caractéristique courant-

tension de la jonction *p-n* en cette polarisation *courbe caractéristique inverse*.

Comme il est visible sur la figure 11.4, *c*, la hauteur de la barrière de potentiel augmente en polarisation inverse de $-eV$.

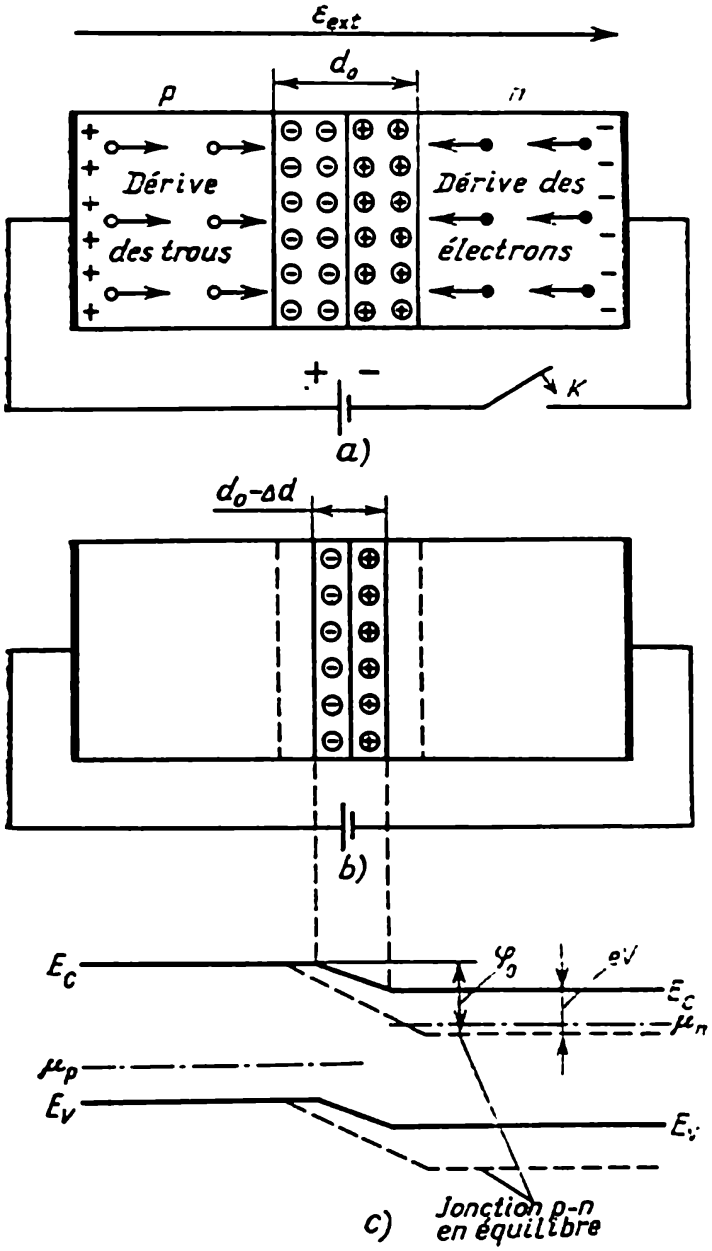

Fig. 11.5. Processus transitoires (*a*) et état stationnaire (*b*, *c*) d'une diode polarisée en sens direct.

En polarisation *directe* (le pôle positif de la source de tension est relié à la région *p* et le pôle négatif à la région *n*; fig. 11.5, *a*) le champ électrique qui prend naissance à l'intérieur des régions *n* et *p* provoque un afflux de porteurs majoritaires vers la région de char-

ge d'espace de la jonction *p-n*. Ce processus continue tant que la différence de potentiel de contact ne tombe à la valeur de $V_c - V$ dans la jonction *p-n*. Dans ces conditions les charges qui ont été initialement créées par la source de f.é.m. extérieure sur les contacts ohmiques se trouvent transportées sur les frontières de la région de charge d'espace et celle-ci se rétrécit jusqu'à $d_0 - \Delta d$. Comme le montre la figure 11.5, *c*, la hauteur de la barrière de potentiel en polarisation directe diminue de eV.

Ainsi, quand on applique à la diode une différence de potentiel extérieure, le circuit extérieur est parcouru, au premier instant, par une impulsion de courant qui assure finalement une augmentation ou une diminution des charges d'espace près des frontières de la jonction *p-n*. C'est pourquoi la jonction *p-n* se comporte comme une capacité. Cette capacité est appelée *capacité de barrière* (ou *de charge*) parce qu'elle est liée à une variation de la barrière de potentiel entre les régions *p* et *n*. Elle a pour valeur

$$C_b = \frac{dQ}{dV}, \tag{11.20}$$

où dQ est la variation de la charge de la jonction *p-n*; dV la variation de la différence de potentiel agissant sur la jonction. L'examen des figures 11.4, *b* et 11.5, *b* permet de conclure qu'une jonction *p-n* plane est pareille à un condensateur plan. Aussi, la valeur de la capacité de barrière peut-elle être déterminée au moyen de la formule donnant la capacité d'un condensateur plan:

$$C_b = \varepsilon\varepsilon_0 S/d, \tag{11.21}$$

où S est la surface de la jonction *p-n*; ε la permittivité relative du semiconducteur; d l'épaisseur de la couche de charge d'espace jouant le rôle de la distance entre les armatures du condensateur.

Une différence par rapport à un condensateur ordinaire tient à ce que d figurant dans l'expression (11.21) n'est pas une grandeur constante mais varie suivant la polarisation extérieure:

$$d = \left[\frac{2\varepsilon\varepsilon_0(\varphi_0 - eV)}{e^2}\,\frac{N_a + N_d}{N_a N_d}\right]^{1/2} = \left[\frac{2\varepsilon\varepsilon_0(\varphi_0 - eV)}{e^2}\,\frac{n_{n0} + p_{p0}}{n_{n0}p_{p0}}\right]^{1/2}. \tag{11.22}$$

Ainsi, la capacité de barrière de la jonction *p-n* est non linéaire. Cela signifie en particulier que l'expression (11.21) n'est applicable qu'à de faibles signaux qui ne font pas varier de façon notable la valeur de d. Un tel signal peut être superposé notamment à la polarisation extérieure continue V. L'expression (11.22) prend une forme beaucoup plus simple pour une jonction *p-n* asymétrique dont une des régions est dopée beaucoup plus fortement que l'autre. Soit par exemple $N_a \gg N_d$, alors

$$d \approx \left[\frac{2\varepsilon\varepsilon_0(\varphi_0 - eV)}{e^2 N_d}\right]^{1/2}. \tag{11.23}$$

En portant (11.23) dans (11.21), on obtient

$$C_b = S\left[\frac{\varepsilon\varepsilon_0 e^2 N_d}{2(\varphi_0 - eV)}\right]^{1/2}. \qquad (11.24)$$

Cette expression est valable pour une jonction *p-n* abrupte dans laquelle les concentrations des impuretés donatrice et acceptrice varient par saut.

Pour des jonctions *p-n* graduelles (diffusées) la capacité de barrière a pour expression

$$C_b = S\left[\frac{(\varepsilon\varepsilon_0)^2 e^2 a}{12(\varphi_0 - eV)}\right]^{1/3}. \qquad (11.25)$$

§ 11.3. Courants direct et inverse dans une jonction *p-n*

Quand le processus de charge ou de décharge de la capacité de barrière prend fin, le courant à travers la diode ne s'annule pas parce que la variation de la hauteur de la barrière de potentiel pour les porteurs majoritaires, lorsque la jonction *p-n* est soumise à une différence de potentiel extérieure, trouble l'égalité des flux de porteurs opposés traversant la jonction et fait naître un courant dans cette dernière. L'intensité de ce courant est facile à calculer dans le cas d'une polarisation inverse suffisamment forte ($-V \gg 4kT/e \approx \approx 0{,}1$ V) quand les flux de porteurs majoritaires à travers la jonction *p-n* pratiquement disparaissent et il n'existe que des flux de porteurs minoritaires qui n'ont à franchir aucune barrière.

Les porteurs minoritaires, par exemple les électrons dans la région *p*, prennent naissance par suite de la génération thermique de paires électron-trou. Grâce à l'agitation thermique désordonnée, les porteurs minoritaires arrivent, pendant la durée de leur vie, à s'éloigner du point où ils « sont nés » à une distance égale en moyenne à la longueur de diffusion $L = \sqrt{D\tau}$, où D est le coefficient de diffusion. (En réalité le chemin parcouru par un porteur pendant sa durée de vie est supérieur de beaucoup d'ordres de grandeur à L.) Si la distance entre la frontière de la couche de charge d'espace et le point de génération d'un porteur est plus petite que la longueur de diffusion, la probabilité pour que ce porteur pénètre, au moins une fois, dans la couche de charge d'espace et soit entraîné par son champ est voisine de l'unité. De tels porteurs constituent un *courant inverse* de la jonction *p-n*. Bien entendu, une partie des porteurs peuvent dans ce cas se recombiner sans atteindre la jonction *p-n*, mais en revanche une partie des porteurs nés à des distances supérieures à L peut atteindre cette jonction. Aussi, en calculant l'intensité d'un courant inverse peut-on considérer la génération de porteurs seulement dans une couche d'épaisseur L, mais estimer qu'ils atteignent tous la jonction *p-n*.

La vitesse de génération thermique g_0 (le nombre de porteurs produits par unité de temps et de volume) peut être déterminée en utilisant son égalité à la vitesse de recombinaison R_0 à l'état d'équilibre thermique. Quant à cette dernière, elle est égale à la concentration des porteurs divisée par leur durée de vie:

$$g_0 = R_0 = n_{p0}/\tau_n \text{ (dans la région } p\text{)}, \qquad g_0 = p_{n0}/\tau_p \text{ (dans la région } n\text{)}. \tag{11.26}$$

En multipliant ces grandeurs par les volumes des parties du semi-conducteur qui apportent une contribution au courant inverse, c'est-à-dire par SL_n et SL_p (S étant la surface de la jonction p-n), et par la charge de l'électron e, on obtient le courant de porteurs minoritaires à travers la jonction p-n:

$$-I_s = -I_{ns} - I_{ps} = eS\left(\frac{n_{p0}}{\tau_n}L_n + \frac{p_{n0}}{\tau_p}L_p\right) = = eS\left(\frac{n_{p0}\sqrt{D_n}}{\sqrt{\tau_n}} + \frac{p_{n0}\sqrt{D_p}}{\sqrt{\tau_p}}\right). \tag{11.27}$$

Le signe moins signifie que le courant va de la région n à la région p suivant la règle des signes adoptée plus haut.

Les porteurs *minoritaires* que le champ électrique fait passer à travers la jonction p-n deviennent, dans des régions correspondantes, porteurs *majoritaires*. La charge qui leur est liée se résorbe d'une manière pratiquement instantanée, comme il arrive toujours à la charge de porteurs majoritaires, et disparaît parce que les porteurs s'en vont dans le circuit extérieur. De (11.27) il résulte que si les polarisations sont suffisamment fortes, le courant à travers la jonction ne dépend pas de V, en gardant une valeur limite I_s. C'est pour cette raison que I_s est appelé *courant de saturation*. En cherchant à exprimer sa nature physique, on lui donne parfois le nom de *courant de thermogénération*.

A l'état d'équilibre (pour $V = 0$) et pour une polarisation directe ($V > 0$) appliquée à la jonction p-n, il se forme dans le semi-conducteur, sous l'effet de la chaleur, un même nombre de porteurs que celui produit dans le cas où la jonction est soumise à une polarisation inverse. Avec la même probabilité ces porteurs peuvent atteindre la jonction p-n et être entraînés par son champ. Cela explique pourquoi la jonction p-n est toujours parcourue par le courant de thermogénération $-I_s$. Dans les conditions d'équilibre ce courant est compensé par le courant opposé de porteurs majoritaires $+I_s$ de sorte que le courant résultant $-I_s + I_s = 0$.

Mais, comme le montre l'expression (11.2), les flux de porteurs majoritaires dépendent de la hauteur de la barrière de potentiel φ. Pour une barrière d'équilibre φ_0 le flux de porteurs majoritaires est

proportionnel à $e^{-\varphi_0/(kT)}$; si la jonction est soumise à une polarisation directe V, la hauteur de la barrière diminue jusqu'à la valeur de $\varphi_0 - eV$ et le flux de porteurs majoritaires devient $e^{eV/(kT)}$ fois plus intense. Le courant de porteurs majoritaires devient égal à $I_{maj} = I_s e^{eV/(kT)}$. Le courant résultant a pour valeur

$$I = -I_s + I_s e^{eV/(kT)} = I_s (e^{eV/(kT)} - 1). \qquad (11.28)$$

Il convient de souligner que le courant produit par les porteurs majoritaires est beaucoup plus faible que leur flux multiplié par la charge de l'électron ($^1/_4 Sen_{n0} \langle v_n \rangle$ et $^1/_4 Sep_{p0} \langle v_p \rangle$) parce que la

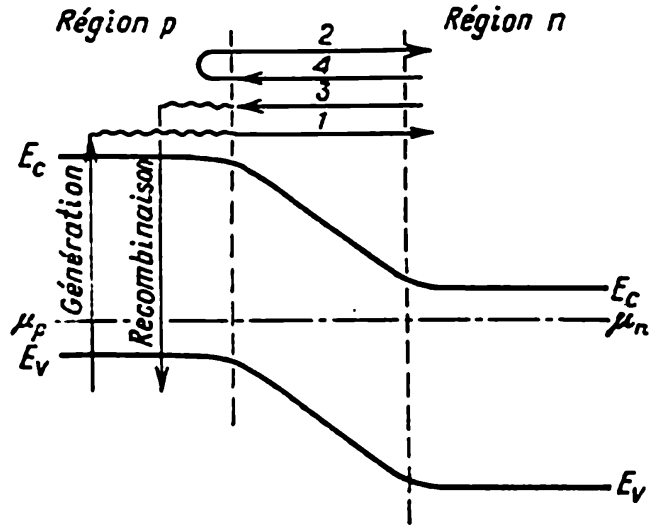

Fig. 11.6. Flux d'électrons à travers une jonction p-n en équilibre.

contribution au courant direct n'est apportée que par ceux des porteurs majoritaires qui se recombinent, après avoir traversé la jonction p-n, dans des régions correspondantes, *sans revenir dans la région d'où ils sont partis*. Or, la partie principale des porteurs arrive justement pendant la durée de leur vie à traverser de nouveau la frontière de la jonction p-n et, étant entraînés par son champ, *reviennent dans leurs régions* (les électrons dans la région n et les trous dans la région p).

La figure 11.6 représente schématiquement les flux d'électrons à travers une jonction p-n à l'état d'équilibre. De la région p s'écoule un flux d'électrons *1* qui sont nés dans cette région par suite de la génération thermique et ont diffusé jusqu'à la région de charge d'espace. Un même flux d'électrons *3* qui se recombinent dans la région p vient à la rencontre du flux *1*. En outre la jonction p-n est traversée par un flux d'électrons *4* venant de la région n et revenant ensuite dans cette région sous la forme du flux *2* qui prend naissance du fait qu'après avoir subi dans la région p une série de collisions, les

électrons pénètrent par hasard dans le champ de charge d'espace qui les repousse dans la région n.

Comme il résulte de (11.28) la forme de la caractéristique courant-tension (CCT) d'une jonction p-n est la même que celle de la CCT d'une barrière de Schottky (v. fig. 10.14) bien que les expressions pour le courant I_s sont dans ces cas nettement différentes.

L'expression que nous avons obtenue pour la caractéristique courant-tension d'une diode à semiconducteur est valable pour des polarisations directes pas trop fortes, quand la barrière de potentiel subsiste encore ($eV < \varphi_0$) et la résistance de la jonction p-n est beaucoup plus grande que celle des régions n et p de la diode adjacentes à la jonction. Lorsque la jonction p-n est polarisée en sens inverse, l'expression (11.28) est applicable pour des tensions inférieures à celles de claquage. En outre, pour établir les expressions (11.27) et (11.28) nous avons négligé la génération thermique et la recombinaison des porteurs à l'intérieur de la couche de charge d'espace elle-même en la considérant comme étroite. Enfin, pour les applications pratiques de l'expression (11.28) il convient d'avoir en vue que la température T qu'elle fait intervenir est celle de la jonction p-n et des régions adjacentes qui peut différer fortement au cours du fonctionnement de la diode de la température ambiante.

Le courant de saturation I_s figurant dans les expressions (11.27) et (11.28) dépend très fortement de T parce que la vitesse de génération thermique croît brutalement lorsque la température augmente. En effet, suivant la loi d'action de masse

$$n_{p0} = \frac{n_i^2}{p_{p0}} = \frac{n_i^2}{N_a}\,; \quad p_{n0} = \frac{n_i^2}{n_{n0}} = \frac{n_i^2}{N_d}\,; \quad n_i^2 \sim e^{-E_g/(kT)}.$$

Aussi, lorsque la température s'élève, le courant inverse croît brutalement et les propriétés redresseuses de la jonction p-n s'en trouvent compromises. Lorsque la température augmente, c'est aussi la hauteur d'équilibre de la barrière de potentiel $\varphi_0 = kT \ln (p_{p0}/p_{n0}) = kT \ln (n_{n0}/n_{p0})$ qui diminue. A la température pour laquelle $n_{n0} \approx n_{p0} \approx n_i$ la jonction disparaît totalement.

§ 11.4. Propriétés d'une jonction p-n lors du fonctionnement en impulsions et en hautes fréquences

Propriétés en impulsions. Un domaine d'emploi bien large des diodes à semiconducteurs est représenté par des circuits impulsionnels utilisés en radio-électronique, dans la technique de calcul électronique et les dispositifs d'automatisme. Le paramètre principal qui détermine l'applicabilité des diodes à cette fin est la *rapidité de fonctionnement* (ou la *rapidité de réponse*) qui caractérise le temps mis par la jonction p-n pour passer de la polarisation directe à la polarisation inverse et inversement de la polarisation inverse à la polarisation directe.

Regardons quels sont les processus, se déroulant dans une jonction *p-n*, qui déterminent ce paramètre. La figure 11.7, *a* représente schématiquement la répartition des porteurs majoritaires et minoritaires dans les régions *p* et *n* d'un semiconducteur lorsque la jonction *p-n* est en équilibre. Lorsqu'on applique à la diode une polarisation directe V, la barrière de potentiel de la jonction s'abaisse de eV et le flux de porteurs majoritaires à travers la jonction *p-n* devient exp $[eV/(kT)]$ fois plus intense, de sorte que les concentrations des trous près de la frontière *2* de la région *n* et des électrons près de la frontière *1* de la région *p* (fig. 11.7, *b*) croissent jusqu'à

$$p_n(0) = p_{n0} \exp [eV/(kT)];$$
$$n_p(0) = n_{p0} \exp [eV/(kT)].$$

Les trous qui sont passés dans la région *n* et les électrons qui sont passés dans la région *p* deviennent dans ces régions porteurs minoritaires.

Ainsi, tout se passe comme si sous l'action de la polarisation directe les porteurs minoritaires étaient « injectés » à travers les frontières de la jonction *p-n* dans des régions correspondantes du semiconducteur. Ce phénomène a reçu le nom d'*injection de porteurs minoritaires*.

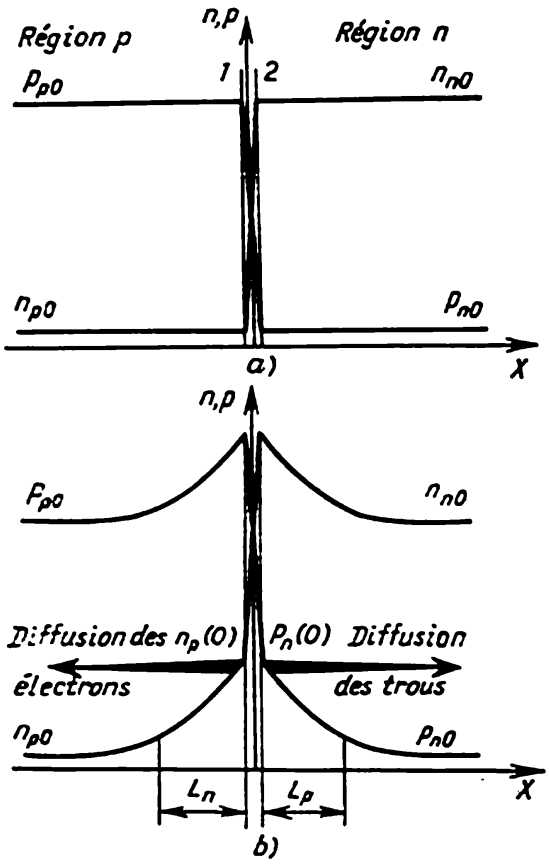

Fig. 11.7. Injection de porteurs de charge minoritaires dans une jonction *p-n* polarisée en sens direct.

Les trous injectés dans la région *n* attirent les électrons se trouvant à l'intérieur de cette région de sorte que la concentration des électrons près de la jonction *p-n* s'élève par rapport à la concentration à l'intérieur (fig. 11.7, *b*). La charge négative des électrons attirés *met sous écran* la charge positive des trous en excès. Une situation analogue se présente aussi dans la région *p*: la charge positive des trous attirés met sous écran la charge négative des électrons injectés. C'est pourquoi les trous en excès et les électrons en excès, injectés respectivement dans les régions *n* et *p*, n'y produisent pas de charges d'espace non compensées dont le champ pourrait empêcher le mouvement des porteurs minoritaires vers l'intérieur du semiconducteur. La pénétration de ces porteurs en profondeur du semiconducteur se fait exclusivement par diffusion dont la vitesse est proportionnelle

au gradient de concentration des trous dp_n/dx dans la région n et au gradient de concentration des électrons dn_p/dx dans la région p.

La figure 11.8, *a* montre la variation avec le temps de la répartition des trous injectés dans la région n. A l'instant initial, après l'application de la tension de polarisation directe (instant t_1 sur la figure 11.8, *b*), le gradient de concentration des trous près de la frontière *2* est très grand (courbe *1* de la figure 11.8, *a*) parce que les trous injectés dans la région n se concentrent dans une couche mince près de cette frontière. C'est pourquoi le courant direct de la diode est dans ce cas bien intense; il n'est limité pratiquement que par la résistance r de ses régions passives (palier *1* de la figure 11.8, *c*). Mais au fur et à mesure de la diffusion des trous dans la région n et de leur recombinaison le gradient de concentration diminue (courbe *2* de la figure 11.8, *a*), ce qui conditionne une diminution du courant direct lui aussi (fig. 11.8, *c*). Pendant un temps approximativement égal à la durée de vie τ_p des trous excédentaires il s'établit leur répartition stationnaire (invariable avec le temps) dans la région n (courbe *3* de la figure 11.8, *a*) et le courant direct atteint sa valeur normale.

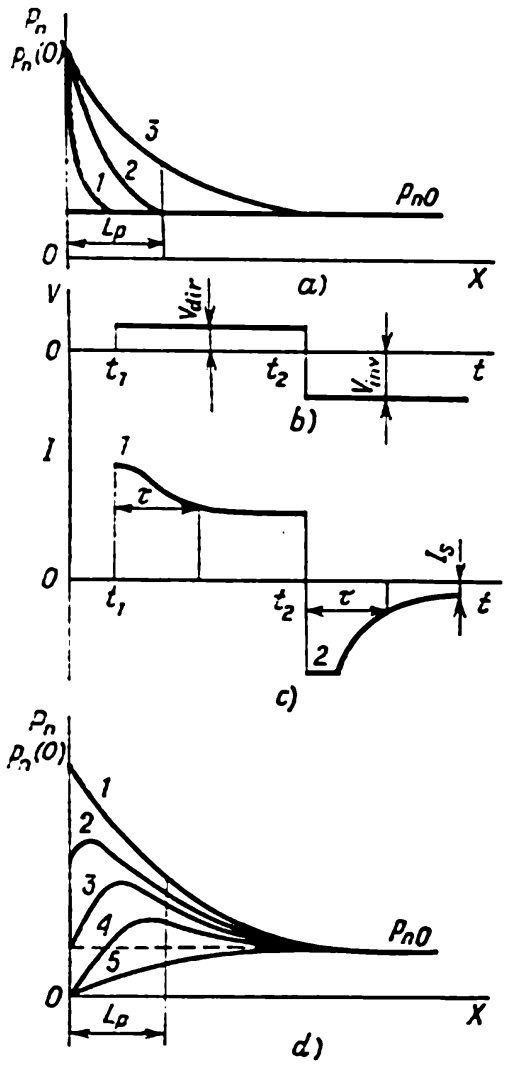

Fig. 11.8. Explication des propriétés impulsionnelles d'une jonction *p-n*.

Des phénomènes analogues se produisent aussi dans la région p où sont injectés des électrons: les processus transitoires et l'établissement d'un courant direct stationnaire se réalisent ici pendant un temps approximativement égal à τ_n.

Si la diode est commutée d'une manière instantanée du sens direct au sens inverse (instant t_2 sur la figure 11.8, *b*), le courant inverse sera au premier instant très élevé parce que la concentration des porteurs minoritaires sur les surfaces frontières *2* et *1* est forte (courbe *1* de la figure 11.8, *d*); l'intensité de ce courant n'est en fait limitée que par la résistance r des régions passives de la diode (palier *2*

de la figure 11.8, *c*). Avec le temps, les porteurs minoritaires excédentaires concentrés près des frontières *1* et *2* se résorbent progressivement grâce à leur passage dans des régions correspondantes du semiconducteur d'une part et à leur recombinaison d'autre part. Le gradient de concentration qui conditionne leur diffusion vers ces frontières diminue (courbes *2*, *3*, *4* de la figure 11.8, *d*) et le courant inverse diminue (fig. 11.8, *c*). Pendant un temps approximativement égal à τ_p pour les trous et à τ_n pour les électrons, il s'établit une répartition stationnaire des porteurs minoritaires près des frontières de la jonction *p-n* (courbe *5* de la figure 11.8, *d*) et le courant inverse atteint sa valeur normale I_s (fig. 11.8, *c*).

Ainsi, en cas de la commutation de la diode elle devient le siège des processus transitoires (accumulation des porteurs minoritaires en polarisation directe et leur résorption en polarisation inverse) qui limitent précisément sa rapidité de fonctionnement. Etant donné que ces processus s'achèvent en général en un temps égal à la durée de vie des porteurs minoritaires excédentaires τ, c'est cette dernière qui détermine la rapidité de fonctionnement des diodes impulsionnelles.

Capacité de diffusion d'une jonction *p-n*. De ce qui précède on peut conclure que par rapport à la commutation la diode se comporte comme une résistance R produite par la région de charge d'espace et shuntée par une capacité C_{dif} due à l'accumulation de la charge de porteurs minoritaires lorsque la diode est polarisée en sens direct et à la résorption de cette charge lorsque la diode est polarisée en sens inverse. Cette capacité porte le nom de *capacité de diffusion de la jonction p-n*. A l'application de la polarisation directe, le courant dans la diode est à l'instant initial principalement un courant de charge de la capacité C_{dif} et son intensité peut être grande. Lorsque la diode est commutée sur le sens inverse, le courant inverse représente à l'instant initial principalement un courant de décharge de la capacité C_{dif} et son intensité peut être elle aussi grande.

Comme le montre le calcul, pour un signal variable de faible amplitude

$$C_{dif} = \frac{e}{2kT}[(I_p + I_{ps})\tau_p + (I_n + I_{ns})\tau_n], \qquad (11.29)$$

où comme il résulte de (11.28), $I_p + I_{ps} = I_{ps} \exp[eV/(kT)]$ et $I_n + I_{ns} = I_{ns} \exp[eV/(kT)]$ (ici, V est la polarisation continue de la jonction *p-n*). Comme $\exp[eV/(kT)] \simeq 0$ même pour des polarisations inverses peu importantes, la capacité de diffusion elle aussi $C_{dif} = 0$. Lorsque la diode est polarisée en sens direct, $I_p \gg I_{ps}$ et $I_n \gg I_{ns}$ et donc

$$C_{dif} \approx \frac{e}{2kT}(I_p\tau_p + I_n\tau_n). \qquad (11.30)$$

Pour une jonction p-n fortement asymétrique avec $N_a \gg N_d$ ($p_{p0} \gg n_{n0}$), les courants totaux I et I_s traversant la jonction sont pratiquement égaux aux courants I_p et I_{ps} et de ce fait,

$$C_{dif} \approx \frac{e}{2kT}(I_p + I_{ps})\tau_p \approx \frac{e}{2kT}(I + I_s)\tau_p. \qquad (11.31)$$

Il résulte des expressions (11.29) à (11.31) que pour réduire la capacité C_{dif} et donc améliorer la rapidité de fonctionnement des diodes, il est nécessaire de réduire la durée de vie τ des porteurs minoritaires en excès, en dopant les régions n et p avec une impureté créatrice de centres de recombinaison efficaces. Une telle impureté pour le germanium et le silicium est notamment l'or: le dopage avec de l'or permet de réduire τ à quelques nanosecondes.

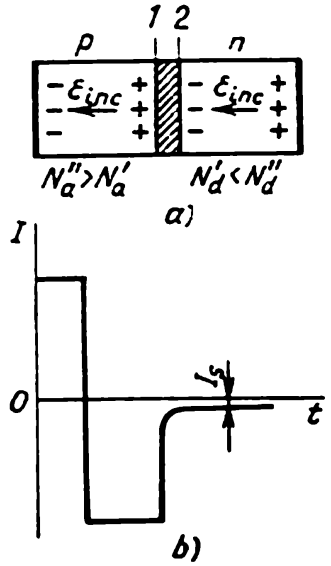

Fig. 11.9. Diodes à champ électrique « incorporé »:

a — schéma de la formation d'un champ « incorporé » par une répartition inégale de l'impureté dopante dans la diode; *b* — caractère de rétablissement du courant inverse de la diode par le champ « incorporé » lors de la commutation de la polarisation directe sur la polarisation inverse.

Un autre procédé permettant d'accroître la rapidité de fonctionnement des diodes consiste à produire dans les régions adjacentes à la charge d'espace de la jonction un champ électrique « incorporé » qui « presse » les porteurs minoritaires vers la jonction p-n. Un tel champ peut être engendré en dopant les régions p et n de telle sorte que la concentration de l'impureté dopante non compensée augmente à mesure qu'on s'éloigne de la jonction p-n (v. fig. 11.1, *a* et 11.3, *a*). Dans ce cas, la concentration des électrons dans la région n ($n_{n0} \simeq N_d$) loin de la jonction sera plus forte que près de la jonction elle-même. De ce fait, les électrons diffuseront vers la jonction en y augmentant la concentration et en la diminuant dans des régions éloignées de la jonction. Le passage des électrons vers la jonction continuera tant que le champ « incorporé » engendré $\mathscr{E}_{inc}$ (fig. 11.9, *a*) n'égalise leurs flux opposés.

Un processus analogue se déroule aussi dans la région p avec les trous, ce qui conduit à la formation d'un champ « incorporé » $\mathscr{E}_{inc}$ aussi dans cette région.

L'examen de la figure 11.9, *a* montre que les champs « incorporés » s'opposent à la diffusion libre des porteurs minoritaires, injectés à travers la jonction p-n (des trous en profondeur de la région n

et des électrons en profondeur de la région *p*) en les repoussant vers les frontières de la jonction *p-n*. De ce fait, la résorption des porteurs minoritaires, lorsque la diode est commutée de la polarisation directe à la polarisation inverse, se termine d'une façon très brusque (fig. 11.9, *b*). Cela permet d'utiliser le top de courant inverse pour la formation d'impulsions à front très raide ($\approx 10^{-12}$ s). De telles diodes ont reçu le nom de *diodes à recouvrement* (ou *à récupération*) *brusque de la résistance inverse* ou *de diodes à accumulation de charges* (DAC).

Enfin, on peut tout simplement se débarrasser de la capacité de diffusion en utilisant par exemple un contact redresseur métal-semiconducteur de type *n*. La caractéristique courant-tension d'un tel contact est la même que celle d'une jonction *p-n* (v. § 10.4), mais on n'y fait pas recours à l'injection de porteurs minoritaires à travers la jonction. L'application de la polarisation directe abaisse la hauteur de la barrière de potentiel pour les électrons passant du semiconducteur dans le métal (v. fig. 10.13, *c*) de sorte que le flux d'électrons du semiconducteur dans le métal augmente brusquement et le contact est parcouru par un courant direct. En polarisation inverse, le courant de saturation est dû aux électrons passant du métal dans le semiconducteur de type *n* (v. fig. 10.13, *b*). Dans les deux cas, les électrons en excès, ayant traversé le contact, « se résorbent » d'une manière instantanée (en un temps de $\approx 10^{-12}$ s) parce que la charge qu'ils ont apportée produit un champ qui regroupe, pratiquement en un instant, les électrons du métal ou du semiconducteur de type *n*.

Les diodes à contact redresseur métal-semiconducteur sont appelées *diodes Schottky*. Dans ces diodes on arrive à réduire le temps de commutation à 10^{-11} s. Ce temps se détermine par le temps de recharge de la capacité de barrière. En principe, on doit tenir compte de la capacité de barrière aussi dans la commutation de la jonction *p-n* mais lorsque la diode fonctionne en impulsions de courant direct, la capacité de barrière est très inférieure à la capacité de diffusion.

Propriétés d'une jonction *p-n* en hautes fréquences. Considérons maintenant le fonctionnement de la diode dans le cas où un signal variable de faible amplitude est superposé à la tension de polarisation continue V. Dans ce cas, la diode peut être représentée par un schéma équivalent simplifié suivant : en parallèle avec une résistance active non linéaire R_a de la joncton *p-n* sont branchées une capacité de diffusion non linéaire C_{dif} et une capacité de barrière C_b ; en série avec ce circuit est mise la résistance r des régions passives de la diode (fig. 11.10). Pour mettre en évidence les particularités de fonctionnement de la diode en hautes fréquences, analysons de plus près ce schéma équivalent.

La résistance active que la jonction *p-n* offre à un faible signal variable de basse fréquence se détermine par la relation

$$R_a = \frac{1}{\mathrm{d}I/\mathrm{d}V} = \frac{kT}{e}\frac{1}{I+I_s}. \qquad (11.32)$$

Par basse fréquence on entend ici une fréquence angulaire ω du signal pour laquelle la période d'oscillations est très supérieure à la durée de vie des porteurs injectés : $1/\omega \gg \tau_p$. Dans ce cas, tous les processus transitoires dans la jonction arrivent à se produire durant un temps $\sim 1/\omega$ si bien que la capacité de diffusion se décrit par la relation (11.31) et la capacité de barrière par la relation (11.24).

A mesure que la fréquence ω s'élève, le temps $1/\omega$ se trouve de plus en plus insuffisant à l'achèvement des processus transitoires. Il doit en résulter une diminution du nombre de porteurs injectés pendant l'alternance positive du signal et donc une diminution de la capacité de diffusion. En outre, les porteurs injectés n'arrivent pas à diffuser dans la profondeur des régions passives de la diode à la longueur de diffusion, en se concentrant, avec un grand gradient, dans une couche mince près des frontières de la jonction *p-n*, ce qui doit conduire à une augmentation du courant direct, c'est-à-dire à une diminution de la résistance active de la jonction *p-n*.

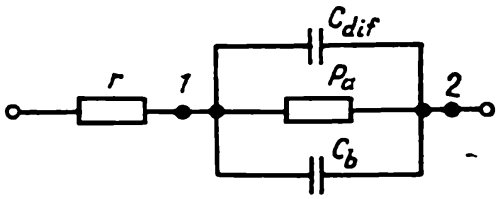

Fig. 11.10. Schéma équivalent d'une diode.

Pendant l'alternance négative du signal, du fait même de la forte concentration des porteurs minoritaires qui n'ont pas réussi à diffuser dans la profondeur des régions passives et se sont localisés dans une couche mince près des frontières de la jonction *p-n*, le courant inverse doit augmenter et donc la résistance active de la jonction polarisée en sens inverse doit diminuer.

Ainsi, l'élévation de la fréquence du signal appliqué à une jonction *p-n* doit conduire à une diminution de la résistance active R_a et de la capacité de diffusion C_{dif}. Comme le montre un calcul, pour une jonction *p-n* asymétrique ($N_a \gg N_d$)

$$R_a = \frac{kT}{e} \frac{\sqrt{2}}{I+I_s} \frac{1}{\sqrt{1+(1+\omega^2\tau_p^2)^{1/2}}}, \tag{11.33}$$

$$C_{dif} = \frac{e(I+I_s)}{\sqrt{2}\,kT} \frac{\tau_p}{\sqrt{1+(1+\omega^2\tau_p^2)^{1/2}}}. \tag{11.34}$$

Pour les basses fréquences ($\omega \ll 1/\tau_p$) ces expressions se réduisent à (11.32) et (11.31). Pour les hautes fréquences ($\omega \gg 1/\tau_p$), les unités figurant sous le radical des expressions (11.33) et (11.34) peuvent être négligées devant $\omega\tau_p$. Alors, il vient

$$R_a = \frac{kT}{e(I+I_s)} \frac{\sqrt{2}}{\sqrt{\omega\tau_p}}, \tag{11.35}$$

$$C_{dif} = \frac{e(I+I_s)}{\sqrt{2}\,kT} \sqrt{\frac{\tau_p}{\omega}}. \tag{11.36}$$

Les conductances correspondant à la résistance R_a et à la capacité C_{dif} ont pour expressions

$$\frac{1}{R_a} = \frac{e(I+I_s)}{\sqrt{2}\,kT}\sqrt{\omega\tau_p}, \tag{11.37}$$

$$\omega C_{dif} = \frac{e(I+I_s)}{\sqrt{2}\,kT}\sqrt{\omega\tau_p}. \tag{11.38}$$

On voit qu'elles sont égales l'une à l'autre: $1/R_a = \omega C_{dif}$.

Comme $I + I_s = I_s \exp[eV/(kT)]$ et $I_s = L_p p_{n0}/\tau_p = p_{n0}\sqrt{D_p\tau_p}/\tau_p \sim 1/\sqrt{\tau_p}$, la résistance active d'une jonction *p-n* en hautes fréquences ne dépend pas de la durée de vie des porteurs injectés et diminue en raison inverse de $\sqrt{\omega}$:

$$R_a \sim 1/\sqrt{\omega}. \tag{11.39}$$

Le comportement de la réactance de la capacité de diffusion C_{dif} est analogue à celui de la résistance:

$$1/(\omega C_{dif}) \sim 1/\sqrt{\omega}. \tag{11.40}$$

La figure 11.11 représente schématiquement la variation de R_a et de C_{dif} en fonction de la fréquence du signal; la droite horizontale indique la capacité de barrière de la jonction *p-n* dont la valeur est indépendante de ω. Comme il est visible sur cette figure, aux fréquences élevées la capacité de diffusion devient inférieure à la capacité de barrière de sorte que la conductance due à la capacité de barrière ωC_b dépasse la conductance due à la capacité de diffusion et la conductance de la jonction *p-n* égale à cette dernière.

La diminution de la résistance de la jonction *p-n* aux fréquences élevées a pour résultat qu'une partie de plus en plus grande de la tension appliquée à la diode chute non pas dans la jonction mais dans la résistance des régions passives de la diode. Aux fréquences pour lesquelles la résistance de la jonction *p-n* devient beaucoup plus petite que r, le courant de la diode polarisée, aussi bien en sens direct qu'en sens inverse, se détermine par r et non par la résistance de la jonction, ce qui signifie que la diode perd ses propriétés redresseuses. Pour fréquence limite (ω_{lim}) de fonctionnement de la diode on prend une fréquence pour laquelle la résistance équivalente de la jonction *p-n* (entre les points *1* et *2*, v. fig. 11.10) devient égale à la résistance r des régions passives. Vu qu'aux hautes fréquences la résistance équivalente de la jonction *p-n* se détermine par la capacité de barrière $1/(\omega C_b)$, pour le calcul de la fréquence limite ω_{lim} il faut comparer r à la réactance de la capacité de barrière C_b et évaluer ω_{lim} à partir de la relation

$$r = 1/(\omega_{lim} C_b). \tag{11.41}$$

De (11.41) il ressort que pour élever la fréquence limite de fonctionnement de la diode il faut réduire le produit rC_b. La résistance r peut être réduite en améliorant la qualité des contacts ohmiques des régions n et p et en réduisant l'épaisseur de ces régions. Il semblerait ensuite que la fréquence ω_{lim} pourrait être élevée aussi par diminution de la surface S de la jonction p-n vu que cela entraîne une diminution de la capacité C_b. Mais ce procédé ne permet d'élever la fréquence ω_{lim} que dans le cas des diodes à pointe. En effet, pour des

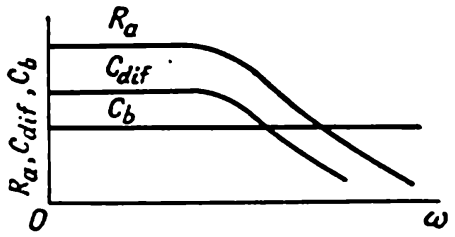

Fig. 11.11. Variations de la résistance, de la capacité de diffusion et de la capacité de barrière d'une jonction p-n en fonction de la fréquence.

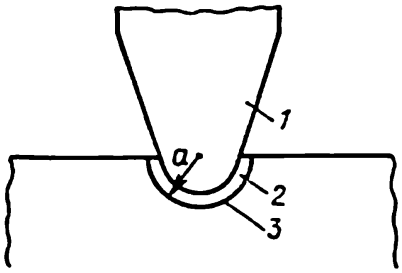

Fig. 11.12. Schéma d'une diode à pointe:
1 — contact métallique; 2 — jonction p-n; 3 — semiconducteur.

jonctions planes $C_b \sim S$ et $r \sim 1/S$ si bien que rC_b ne dépend pas de S. Il n'en est plus de même pour des diodes à pointe dont $r = \rho/(\pi a) \sim 1/\sqrt{S}$, où ρ est la résistivité du semiconducteur; a le rayon d'un contact semi-sphérique; $S = 2\pi a^2$ la surface de la jonction (fig. 11.12). Comme $C_b \sim S$, on a $C_b r \sim \sqrt{S}$ et $\omega_{lim} \sim 1/\sqrt{S}$. Aussi, en diminuant S, peut-on donc élever la fréquence limite ω_{lim}.

La réduction de la résistivité ρ des régions passives de la diode permet elle aussi d'élever ω_{lim}. En effet, pour une jonction p-n asymétrique, dont la région n est à forte résistivité, on a $r \sim 1/n_{n0}$ et $C_b \sim 1/d \sim \sqrt{n_{n0}}$ [v. (11.24)] de sorte que $rC_b \sim 1/\sqrt{n_{n0}} \sim \sqrt{\rho}$, où ρ est la résistivité de la région n. Mais il faut ne pas perdre de vue que lorsque le degré de dopage augmente, la tension de claquage d'une jonction p-n diminue et que finalement on obtient d'abord une diode inverse (dont la fréquence de fonctionnement limite est la plus élevée) et puis, avec un dopage encore plus fort, une diode tunnel qui est inapte à servir de détecteur (v. plus loin).

§ 11.5. Claquage d'une jonction p-n

La figure 11.13 montre comment varie le courant inverse d'une jonction p-n lorsque la tension inverse augmente. Pour une valeur déterminée de $V_{inv} = V_{cl}$, on constate que le courant inverse croît brutalement. Ce phénomène a reçu le nom de *claquage* de la jonction

et la tension V_{cl} à laquelle il se produit est appelée *tension de claquage*.

Suivant le caractère des processus physiques qui conditionnent une croissance brusque du courant inverse, on distingue quatre types principaux de claquage: *par effet tunnel*, *par avalanche*, *thermique* et *superficiel*.

Claquage thermique. Lorsqu'une jonction *p-n* est parcourue par un courant inverse, il s'y dégage une chaleur et sa température s'élève. L'élévation de la température se détermine par la qualité du dispositif d'extraction de la chaleur caractérisée par une résistance thermique R_θ. Cette résistance est égale à l'accroissement de température ΔT de la jonction par une puissance unité W dégagée dans la jonction: $\Delta T = R_\theta W$. L'augmentation de la température provoque une augmentation du courant inverse ce qui, à son tour, fait croître de nouveau la température et le courant inverse et ainsi de suite. Pour une puissance W déterminée, dont la valeur est d'autant plus grande que la résistance thermique R_0 de l'appareil est plus faible, le courant inverse commence à croître en avalanche et il se produit un claquage thermique de la jonction *p-n*. Dans le cas d'un claquage thermique la caractéristique courant-tension peut comporter une partie à résistance différentielle négative quand la croissance du courant s'accompagne d'une diminution de la tension agissant sur la diode (courbe *1* de la figure 11.13).

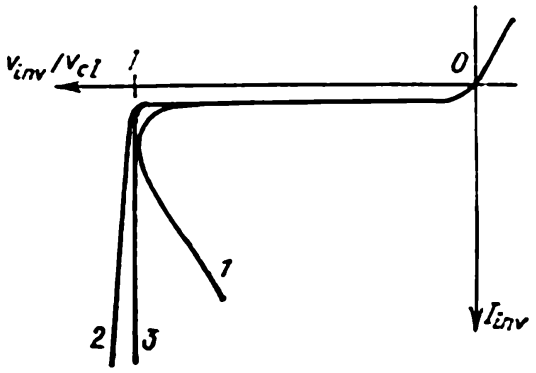

Fig. 11.13. Claquage d'une jonction *p-n*:
1 — thermique; *2* — par effet tunnel; *3* — par avalanche.

Claquage par avalanche. Dans des jonctions *p-n* suffisamment larges, soumises à des tensions inverses élevées, les porteurs minoritaires peuvent acquérir dans le champ de la jonction une énergie cinétique si forte qu'ils deviennent capables de provoquer une ionisation par choc du semiconducteur. Ce processus est représenté schématiquement sur la figure 11.14. L'ionisation par choc peut être produite aussi bien par les électrons (fig. 11.14, *a*) que par les trous (fig. 11.14, *b*). Dans les deux cas, l'électron *1* perd de l'énergie tout en restant dans sa bande d'énergie. Il transfère cette énergie à un électron *2* de la bande de valence en le faisant passer dans la bande de conduction et en créant ainsi une nouvelle paire électron-trou. Dans ce cas le courant inverse peut accuser une croissance par avalanche qui conduit à un claquage par avalanche de la jonction.

Supposons que par exemple un électron de la région *p* de la diode pénètre dans la région de charge d'espace de la jonction *p-n* polarisée en sens inverse (flèche *1* de la figure 11.15). Si cette polarisation

est suffisamment forte, cet électron fait naître, avant de quitter la jonction, une nouvelle paire électron-trou et ensuite passe avec l'électron nouvellement créé de la jonction *p-n* dans la région *n*. Le trou qui apparaît dans ces conditions se déplace dans le champ de la jonction *p-n* vers la région *p* de la diode (flèche *2* de la figure 11.15).

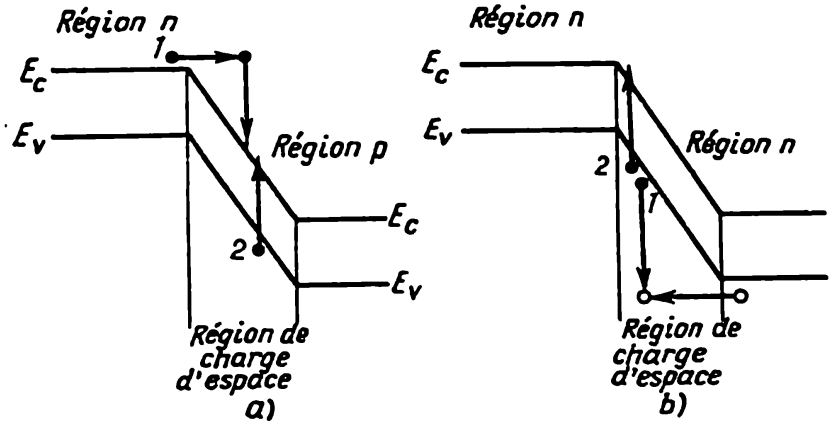

Fig. 11.14. Ionisation par choc dans une jonction *p-n* produite par un électron (*a*) et un trou (*b*).

Avant de quitter la région de charge d'espace ce trou peut former encore une paire électron-trou ; l'électron de cette paire se déplacera vers la région *n* (flèche *3*) et ainsi de suite. Ainsi, au lieu d'un seul électron entrant dans la jonction *p-n* il en sort beaucoup d'électrons

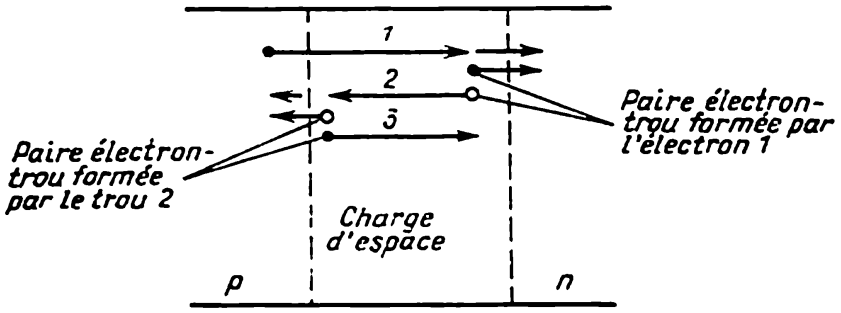

Fig. 11.15. Schéma du développement d'un claquage par avalanche.

et de trous — le courant inverse croît brutalement. Dans la région de claquage, la variation du courant inverse, avec l'augmentation de la tension inverse, est très forte (courbe *3* de la figure 11.13). Cet effet est utilisé pour la stabilisation de tension. Les diodes destinées à servir en un tel régime sont appelées *stabilitrons*. On les fabrique en silicium parce que les diodes à silicium possèdent une branche assez abrupte de la caractéristique courant-tension dans la région de claquage et peuvent fonctionner, sans subir de claquage thermique, dans une large gamme de courants de service.

Les calculs montrent que la tension de claquage par avalanche dépend du degré de dopage des régions p et n de la diode et qu'elle diminue lorsque la concentration de l'impureté dopante augmente. Dans des diodes à forte concentration de l'impureté pour lesquelles le claquage se manifeste pour des tensions inférieures à 4-8 V (pour les diodes au silicium) c'est un claquage d'un autre type, à savoir le claquage par effet tunnel, qui se produit avant le claquage par avalanche (pour des tensions moindres).

Claquage par effet tunnel. Lorsque la tension de polarisation inverse appliquée à une jonction p-n est suffisamment élevée, les

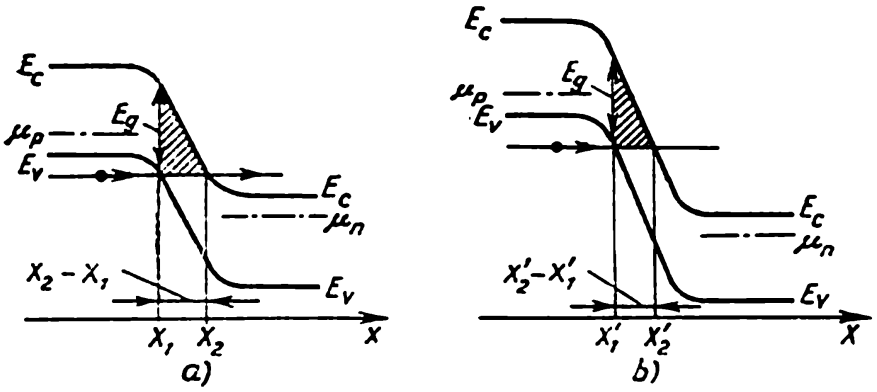

Fig. 11.16. Transitions tunnel dans une diode polarisée faiblement (*a*) et fortement (*b*) en sens inverse.

niveaux d'énergie occupés de la bande de valence de la région p du semiconducteur se situent en face des niveaux non occupés de la bande de conduction de la région n (fig. 11.16, *a*). Dans ce cas, il peut y avoir une transition tunnel directe des électrons de la bande de valence de la région p dans la bande de conduction de la région n. Ces électrons franchissent, par effet tunnel, une barrière de potentiel d'épaisseur $x_2 - x_1$ et de hauteur variant de E_g aux points x_1 à 0 au point x_2. Si V_{inv} augmente, l'épaisseur de la barrière diminue (fig. 11.16, *a*) et l'intensité $\mathscr{E}$ de son champ s'accroît. Il convient de souligner qu'il s'agit non pas de la largeur de la région de charge d'espace (elle s'accroît lorsque la polarisation inverse augmente) mais de la barrière que les électrons franchissent par effet tunnel. Si la jonction p-n est suffisamment mince, le champ $\mathscr{E}$ atteint, même pour une valeur relativement peu élevée de la tension inverse V_{inv}, une valeur pour laquelle on constate un passage intense des électrons à travers la jonction p-n par effet tunnel et le claquage de la jonction. Dans le cas du germanium cela se produit pour $\mathscr{E}_{cl} \approx 3 \cdot 10^7$ V/m et dans le cas du silicium pour $\mathscr{E}_{cl} \approx 10^8$ V/m. On dit que c est un *claquage par effet tunnel.* La courbe caractéristique courant-tension inverse correspondant à ce type de claquage est montrée sur la figure 11.13 (courbe *2*).

Proposons-nous de déterminer les facteurs dont dépend la tension de claquage par effet tunnel. L intensité de champ dans une jonction *p-n* abrupte peut être calculée à partir de (11.14) et (11.15), en dérivant φ par rapport à x:

$$\mathscr{E} = -\frac{dV}{dx} = \frac{1}{e}\frac{d\varphi}{dx}. \tag{11.42}$$

Il n'est pas difficile de s'assurer que le champ $\mathscr{E}$ atteint sa valeur maximale sur la surface frontière entre les régions *p* et *n* où il devient égal à

$$\mathscr{E}_{max} = -\frac{eN_d}{\varepsilon\varepsilon_0} d_n = -\frac{eN_a}{\varepsilon\varepsilon_0} d_p. \tag{11.43}$$

En admettant, pour simplifier, que la jonction *p-n* est asymétrique et que la région *n* est faiblement dopée, en portant $d_n \approx d$ de (11.23) dans (11.43) et en élevant au carré le premier et le second membre de (11.43), on obtient

$$\mathscr{E}^2_{max} = 2\,(\varphi_0 + eV_{inv})/(\varepsilon\varepsilon_0 N_d). \tag{11.44}$$

Ici, au lieu de $-V$ est écrite la grandeur positive V_{inv}. En égalant $\mathscr{E}_{max}$ à $\mathscr{E}_{cl}$, on trouve la tension de claquage par effet tunnel:

$$V_{cl} = ({}^1/_2\,\varepsilon\varepsilon_0 N_d \mathscr{E}^2_{cl} - \varphi_0)/e. \tag{11.45}$$

Comme il résulte de (11.45), la tension de claquage par effet tunnel diminue linéairement lorsque la concentration de l'impureté N_d dans la région faiblement dopée de la diode s'accroît. Pour une certaine concentration le tunneling d'électrons peut commencer même pour $V_{inv} = 0$. La courbe caractéristique courant-tension d'une telle diode est représentée sur la figure 11.17. C'est une *diode* dite *inversée*. Elle possède une caractéristique courant-tension très fortement non linéaire pour des tensions voisines de zéro et est généralement utilisée pour la détection de signaux faibles de haute fréquence.

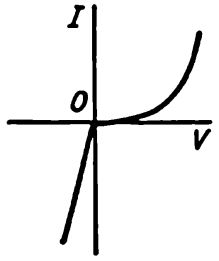

Fig. 11.17. Caractéristique courant-tension d'une diode inversée.

En augmentant encore le degré de dopage, on obtient des diodes dites *tunnel* (ou diodes d'Esaki du nom de leur inventeur).

Claquage superficiel. La charge localisée sur la surface du semiconducteur à l'endroit de sortie de la jonction *p-n* peut provoquer une forte variation de l'intensité du champ dans la jonction et de la largeur de la jonction. Dans ce cas, c'est le claquage de la région adjacente à la surface de la jonction *p-n* qui peut être le plus probable.

§ 11.6. Diodes tunnel

Les diodes tunnel (elles ont été créées en 1958) sont fabriquées avec des semiconducteurs fortement dopés se trouvant à l'état dégénéré. Dans de tels semiconducteurs, les niveaux de Fermi ne se situent pas dans la bande interdite: le niveau de Fermi de la région *p* est descendu dans la bande de valence et celui de la région *n* est monté dans la bande de conduction. Le diagramme énergétique d'une jonction tunnel en équilibre est montré sur la figure 11.18, *a*.

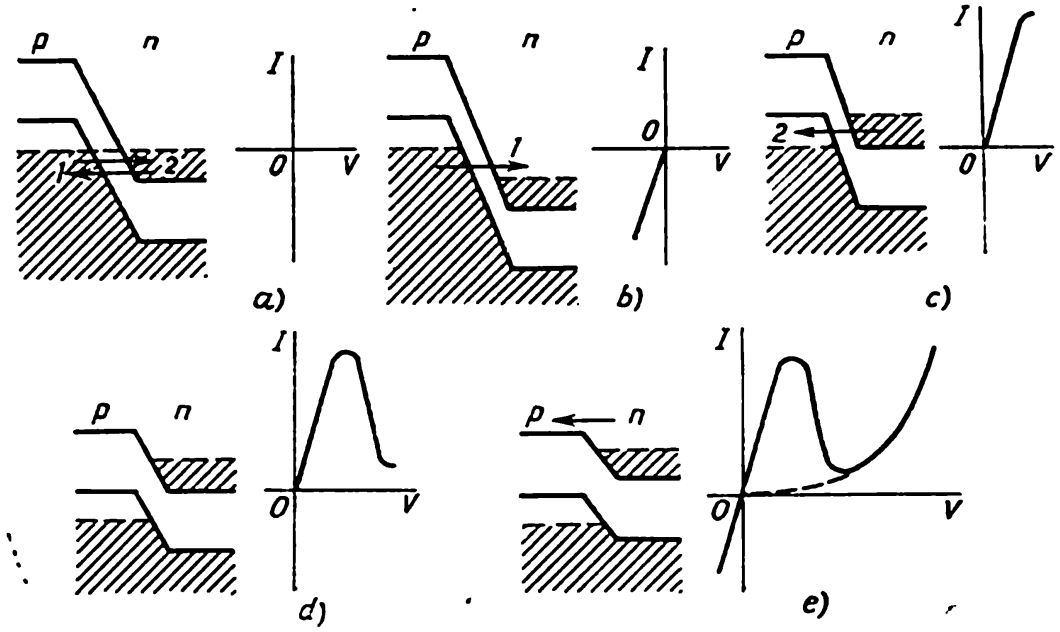

Fig. 11.18. Explication de la caractéristique courant-tension d'une diode tunnel.

Comme il est visible sur la figure 11.18, *a*, il y a recouvrement des bandes: la bande de conduction de la région *n* recouvre partiellement la bande de valence de la région *p*.

Le dopage fort de ces régions a pour effet une faible épaisseur de la jonction *p-n* et une intensité élevée de son champ. Cela rend possible le passage des électrons par effet tunnel de la région *p* dans la région *n* (transitions *1*) et de la région *n* dans la région *p* (transitions *2*). Le flux *1* constitue un courant tunnel inverse et le flux *2* un courant tunnel direct. En l'absence de champ extérieur, ces courants sont égaux de sorte que le courant résultant à travers la jonction est nul.

L'application d'une polarisation inverse à la jonction *p-n* a pour effet d'augmenter le recouvrement des bandes (fig. 11.18, *b*) de sorte qu'en face des états occupés de la bande de valence de la région *p* se trouvent situés des états libres de la bande de conduction de la région *n*. Ceci étant, le nombre des électrons passant de gauche à droite, est plus grand que celui des électrons passant de droite à gauche, de

sorte que dans la jonction il apparaît un courant tunnel inverse dont la valeur augmente en même temps que la polarisation inverse.

Lorsque la jonction est polarisée en sens direct, en face des états occupés du niveau inférieur de la bande de conduction de la région *n* se trouvent situés des états libres du sommet de la bande de valence de la région *p* de sorte qu'un courant tunnel direct prend naissance dans la jonction (fig. 11.18, *c*). Ce courant augmente lorsque la polarisation directe augmente et atteint sa valeur maximale pour une polarisation pour laquelle le niveau inférieur de la bande de conduction de la région *n* se situe au même niveau que le niveau de Fermi de la région *p*. Si la polarisation directe augmente encore, le nombre d'états occupés de la région *n* situés en face des états libres de la région *p* diminue, ce qui provoque une diminution du nombre de passages des électrons de droite à gauche et, comme conséquence, une diminution du courant direct. Le courant tunnel direct tombe à zéro pour une polarisation telle que le niveau inférieur de la bande de conduction de la région *n* se situe à une même hauteur avec le sommet de la bande de valence de la région *p* (fig. 11.18, *d*). Mais pour cette tension il se produit déjà l'injection de porteurs de charge minoritaires, comme dans une diode classique, et un courant de diffusion prend naissance dont l'intensité croît exponentiellement lorsque la polarisation directe augmente (fig. 11.18, *e*).

L'examen de la figure 11.18, *e* montre que la propriété essentielle de la caractéristique courant-tension d'une diode tunnel est de présenter une partie descendante où *la résistance différentielle est négative*. Cette propriété des diodes tunnel permet de les utiliser pour la production d'oscillations électromagnétiques ainsi qu'en qualité de commutateurs, de mélangeurs, etc. L'accumulation et la résorption des porteurs minoritaires n'existant pratiquement pas dans la région des courants tunnel, les diodes tunnel se caractérisent par des réponses fréquentielles et une rapidité de fonctionnement meilleures que celles des diodes classiques.

§ 11.7. Diodes à avalanche

Le phénomène de claquage par avalanche d'une jonction *p-n* est à la base du fonctionnement des *diodes à avalanches* (diodes IMPATT) utilisées comme oscillateurs très haute fréquence. Pour réaliser un tel oscillateur il faut avoir un élément à résistance différentielle négative, c'est-à-dire un élément dans lequel l'augmentation de la tension conduit à une diminution du courant et inversement. La diode tunnel que nous avons examinée plus haut fournit un exemple de tel élément. Pourtant ce sont les diodes à avalanche qui ont trouvé le plus d'applications. Bien que la caractéristique courant-tension stationnaire de la diode à avalanche (v. courbe *3* de la figure 11.13) ne présente pas de partie à pente négative, en courant alternatif de

fréquence déterminée la variation de la tension appliquée à la diode se trouve en opposition de phase avec la variation du courant, si bien que pour un tel signal la diode devient élément à résistance négative.

Montrons comment cela se produit. Il convient de remarquer avant tout que le champ électrique dans une jonction *p-n* étant inhomogène et maximal près de la surface de séparation des régions *p*

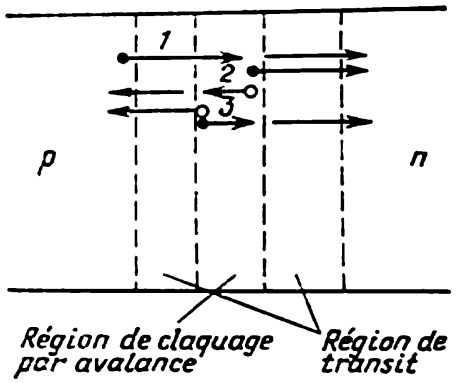

Fig. 11.19. Régions de claquage et de transit lors du claquage par avalanche d'une jonction *p-n*.

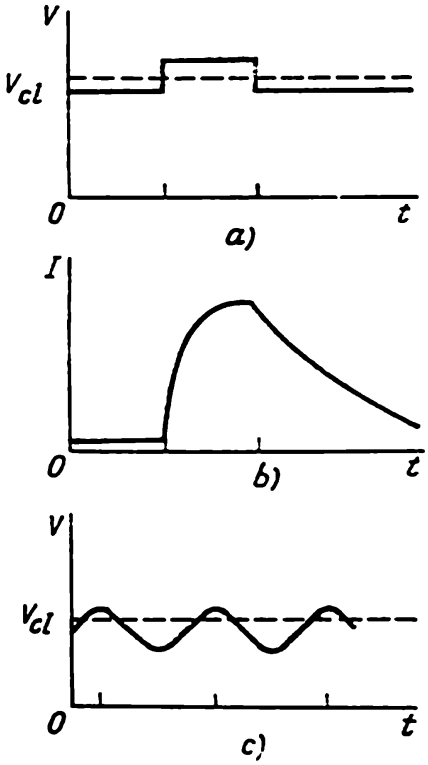

Fig. 11.20. Explication du principe de fonctionnement d'une diode à avalanche à temps de transit (diode IMPATT).

et *n*, c'est précisément là que les électrons et les trous sont « échauffés » assez fortement pour provoquer l'ionisation par choc et donc c'est là que se développe un claquage. Quant à la partie restante de la région de charge d'espace, les porteurs de courant la traversent sans subir d'ionisation. Ces deux régions — la région de claquage par avalanche et la région de transit — de la jonction *p-n* sont représentées sur la figure 11.19 qui est une variante plus précise de la figure 11.15.

Considérons les processus transitoires qui se déroulent dans la diode lorsqu'à sa polarisation continue inverse, légèrement inférieure à la tension de claquage V_{cl}, est superposée une impulsion telle que la tension agissant sur la diode devient supérieure à la tension de claquage (fig. 11.20, *a*). Dans ce cas, le courant à travers la diode sera

limité par la résistance série de la diode et la résistance du circuit extérieur. Au front avant de l'impulsion considérée le courant s'établit pendant le temps de développement de l'avalanche dans la région de claquage (c'est-à-dire dans la partie centrale de la jonction *p-n*) et le temps de charge de la capacité de barrière de la jonction *p-n*. Après la fin de l'impulsion le courant décroît pendant le temps de décharge de la capacité de barrière et, ce qui est l'essentiel, pendant le temps de transit des porteurs de charge, nés dans la région de claquage, dans la région de transit. Le « centre de gravité » de l'impulsion de courant se trouve, somme toute, décalé par rapport à l'impulsion de tension (fig. 11.20, *b*), c'est-à-dire que l'impulsion de courant est retardée sur l'impulsion de tension. Si au lieu du signal impulsionnel, on superpose maintenant à la polarisation continue un signal sinusoïdal (fig. 11.20, *c*), un déphasage se produira entre le premier harmonique de courant et la tension de valeur supérieure à $\pi/2$ si la fréquence du signal est suffisamment élevée. Cela est déjà suffisant pour que la puissance du courant alternatif de fréquence considérée ne soit pas absorbée dans la diode mais se dégage (la quantité $IV \cos \varphi$, où φ est le déphasage entre le courant et la tension, devient négative). La génération d'oscillations de très haute fréquence pour le claquage par avalanche des appareils au germanium a été réalisée pour la première fois en U.R.S.S. en 1959.

Actuellement les diodes à avalanche à temps de transit sont fabriquées principalement au silicium et à l'arséniure de gallium. Elles possèdent le plus souvent une structure plus complexe que celle d'une jonction *p-n* abrupte simple. Elle comporte une région allongée de transit en semiconducteur faiblement dopé, qui adhère à une région de claquage étroite constituée de couches *p* et *n* plus fortement dopées. L'industrie soviétique fabrique des générateurs à diodes à avalanche à temps de transit qui assurent une puissance de sortie près de 1 W dans la gamme de 3 cm de longueur d'onde. Le rendement de conversion de la puissance de courant continu en puissance à très haute fréquence atteint 60 %.

§ 11.8. Phénomènes photo-électriques dans une jonction *p-n*

Effet photovoltaïque. Lorsqu'une jonction *p-n* ou ses régions adjacentes sont soumises à un rayonnement lumineux capable de provoquer la création de paires électron-trou (v. chap. 9) cette jonction est parcourue par un *courant* dit *photo-électrique* ou *photocourant primaire*. Cet effet porte le nom *d'effet photovoltaïque*. Examinons sa nature physique.

La figure 11.21, *a* montre une diode dont la région *p* est éclairée par un flux lumineux de puissance W_0 qui provoque la génération de paires électron-trou. Le nombre G de telles paires qui sont créées

chaque seconde dans le semiconducteur est déterminé par l'expression (9.13). Suivant le rapport qui existe entre la profondeur de pénétration de la lumière dans le semiconducteur $1/\alpha$, la distance w à la couche de charge d'espace de la jonction p-n et l'épaisseur d de cette couche, la lumière est absorbée essentiellement dans la région éclairée de la diode, dans la couche de charge d'espace ou derrière cette couche. Supposons, pour fixer les idées, que la lumière est absorbée dans

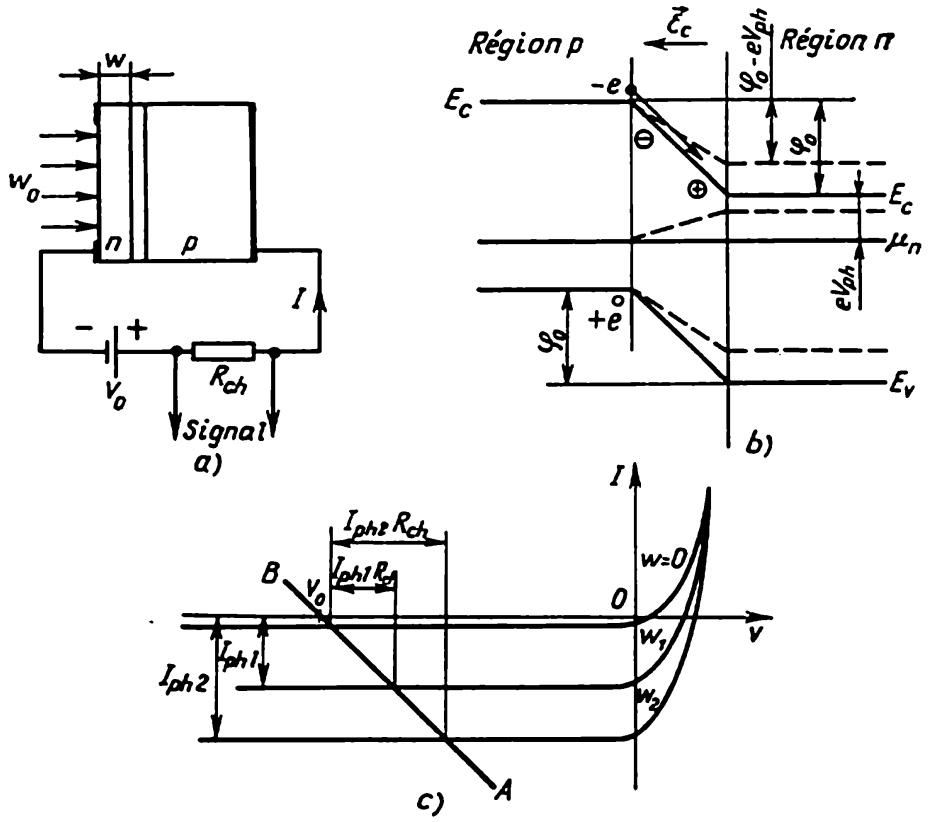

Fig. 11.21. Photorécepteur à jonction p-n:
a — structure et schéma de branchement; *b* — diagramme énergétique d'une cellule photovoltaïque; *c* — caractéristiques courant-tension pour différents niveaux d'éclairement ($W_2 > W_1$).

une couche étroite près de la surface à partir de laquelle les porteurs de charge diffusent ensemble en profondeur du semiconducteur. Si la jonction p-n se situe à une profondeur $w < L$, où L est la longueur de diffusion, une partie considérable des porteurs atteindra la région de charge d'espace de la jonction. Les électrons arrivés vers la jonction p-n sont sollicités par le champ de contact $\mathscr{E}_c$ dirigé de la région n vers la région p, qui les fait passer dans la région n qu'ils chargent négativement (fig. 11.21,*b*). Le photocourant primaire qu'ils conditionnent a pour valeur

$$I_{ph} = e\beta G, \tag{11.46}$$

où β est un coefficient de réunion égal à la part relative des porteurs hors d'équilibre qui atteignent la jonction *p-n* sans se recombiner. Les trous arrivant vers la jonction *p-n* sont incapables de franchir la barrière de potentiel φ_0 et restent donc dans la région *p* en la chargeant positivement. Il en résulte la formation sur la jonction d'une polarisation directe V_{ph} qui réduit la barrière de potentiel à une valeur $\varphi_0 - eV_{ph}$ et fait apparaître un courant direct. Si le circuit extérieur est ouvert, la tension V_{ph} croîtra tant que le courant direct à travers la jonction *p-n* $I_s(e^{eV_{ph}/(kT)} - 1)$ n'égalise pas le photocourant I_{ph}. La différence de potentiel V_{ph} qui prend naissance dans ce cas est appelée *force photo-électromotrice* (fig. 11.21, *b*).

Ainsi, à l'état stationnaire, une jonction *p-n* ouverte est parcourue par un courant direct de porteurs majoritaires $I_s \exp[eV_{ph}/(kT)]$ et, à sa rencontre, par un courant inverse $-I_s$ et un photocourant $-I_{ph}$. Le courant total à travers la jonction *p-n* est nul :

$$\begin{gathered} I_s e^{eV_{ph}/(kT)} - I_s - I_{ph} = 0, \\ I_{ph} = I_s (e^{eV_{ph}/(kT)} - 1). \end{gathered} \tag{11.47}$$

En partant de cette relation il est facile de déterminer la force photo-électromotrice

$$V_{ph} = \frac{kT}{e} \ln \left(\frac{I_{ph}}{I_s} + 1 \right). \tag{11.48}$$

Lorsque la jonction *p-n* est branchée sur une résistance de charge R_{ch}, une partie des porteurs excités par la lumière et transportés à travers la jonction *p-n* revient tandis qu'une autre partie constitue un courant $-I$ passant par le circuit extérieur. Dans ce cas la relation (11.47) doit être récrite sous la forme suivante :

$$\begin{gathered} I_s e^{eV_{ph}/(kT)} - I_{ph} - I_s = -I, \\ I_{ph} = I_s \left(e^{\frac{eV_{ph}}{kT}} - 1 \right) + I. \end{gathered} \tag{11.49}$$

Cette relation permet de déterminer la force photo-électromotrice V_{ph} :

$$V_{ph} = \frac{kT}{e} \ln \left(\frac{I_{ph} - I}{I_s} + 1 \right). \tag{11.50}$$

Si la jonction *p-n* est mise en court-circuit, pratiquement tous les porteurs de charge produits par la lumière cheminent dans le circuit extérieur de sorte que $V_{ph} \approx 0$ et comme il résulte de (11.49), le courant dans le circuit est égal à I_{ph}.

Lorsqu'une jonction *p-n* éclairée est attaquée par une tension de polarisation extérieure V, elle est parcourue par un courant d'obscuri-

té I_s [exp [$eV/(kT)$] — 1] comme une diode ordinaire soumise à une polarisation et par le photocourant primaire $-I_{ph}$ dont la valeur dépend de la puissance W_0 du flux lumineux. Le courant total à travers la jonction a pour valeur

$$I = I_s (e^{eV/(kT)} - 1) - I_{ph}. \qquad (11.51)$$

Les caractéristiques courant-tension d'une jonction *p-n* éclairée sont montrées sur la figure 11.21, *c*. On les obtient en déplaçant les caractéristiques d'une diode non éclairée vers le bas de I_{ph} suivant l'axe des courants.

La répartition de la tension V_0 de la source de polarisation entre la résistance de charge R_{ch} et la diode peut être déterminée soit en portant dans l'expression (11.51), au lieu de V, la quantité $V_0 - IR_{ch}$ et en résolvant numériquement l'équation ainsi obtenue par rapport à I, soit graphiquement à l'aide de la droite de charge AB (fig. 11.21, *c*).

Le phénomène de variation de l'intensité de courant dans le circuit extérieur ou de la tension entre les bornes de la résistance de charge R_{ch} lorsque la jonction *p-n* est éclairée avec un flux lumineux, est utilisé pour l'enregistrement et la mesure de la puissance de ce flux. Les appareils utilisant ce principe jouent le rôle des *photorécepteurs*.

Si une diode est soumise à une polarisation négative qui le reste aussi lorsque la diode est éclairée, on dit qu'elle fonctionne en *régime de photodiode*. Si un photorécepteur est utilisé sans polarisation extérieure, on dit qu'il fonctionne en régime de *photocellule*. Certaines jonctions *p-n* ne peuvent pas fonctionner pour une polarisation inverse quelque peu importante de sorte qu'on ne les utilise que comme photocellules (cellules au sélénium, jonctions semiconductrices à bande interdite étroite et autres).

Cellules photovoltaïques (ou cellules à couche d'arrêt, ou photopiles). La création d'une force photo-électromotrice dans une jonction *p-n* éclairée et le passage d'un courant électrique dans le circuit extérieur permettent d'assurer à l'aide de cellules photo-électriques une conversion directe de l'énergie lumineuse en énergie électrique. Un tel principe est à la base de fonctionnement des piles solaires utilisées pour l'alimentation des équipements radio-électriques à bord des avions et des vaisseaux spatiaux, ainsi que dans des installations productrices d'énergie au sol. La puissance qu'on peut prélever à une cellule photo-électrique a pour valeur

$$W = IV_{ph}.$$

Elle dépend de la résistance de charge R_{ch} et atteint sa valeur maximale pour une certaine valeur de cette résistance. Le rapport de cette puissance à la puissance du rayonnement incident donne le *rende-*

ment de conversion. La valeur de ce rendement dépend du degré de recouvrement de la région de sensibilité spectrale de la cellule photo-électrique et du spectre de lumière solaire, de la résistance interne de la cellule photo-électrique et de certains autres facteurs. La caractéristique lumière-courant d'une cellule photo-électrique exprimant la variation de l'intensité de courant dans le circuit en fonction de la puissance du flux lumineux incident est dans le cas général non linéaire.

Actuellement les matériaux semiconducteurs principaux que l'industrie utilise pour la fabrication de piles solaires sont le silicium et l'arséniure de gallium. Les calculs montrent que le rendement limite (théorique) des convertisseurs au silicium doit être de $\sim$ 20 % et en pratique on a réalisé un rendement de $\sim$ 19 %. Pour un emploi dans les installations à terre il est bien entendu nécessaire de réduire fortement le coût des piles solaires. C'est pourquoi, ces derniers temps de nombreux travaux sont consacrés à la réalisation de piles à semiconducteurs amorphes déposés par pulvérisation sur des bandes d'acier, des verres à vitres, etc.

En plus de l'utilisation des cellules photo-électriques comme convertisseurs de l'énergie solaire en énergie électrique, on les utilise aussi comme capteurs sensibles réagissant à la variation de l'intensité du flux lumineux. A cette fin on utilise très largement des cellules au germanium, à oxyde de cuivre, au sélénium, au sulfure d'argent, au sulfure de thallium et autres. Leur sensibilité intégrale est d'environ 2 à 3 ordres de grandeur supérieure à celle des cellules photo-émissives. Pour améliorer cette sensibilité les cellules sont réalisées de telle sorte que le nombre de porteurs de charge créés par l'éclairement et atteignant la jonction *p-n* soit aussi grand que possible. A cet effet, la base w de la cellule (fig. 11.21, *a*) est rendue aussi mince que possible et le matériau semiconducteur choisi doit avoir une longueur de diffusion L des porteurs aussi grande que possible de façon à satisfaire à la relation $w \ll L$.

Les cellules photovoltaïques ont un inconvénient sérieux de présenter une inertie relativement grande. Cela limite l'emploi de telles cellules en qualité de capteurs de flux lumineux modulés à haute fréquence.

Photodiodes. Lorsqu'un photorécepteur fonctionne en régime de diode, on lui applique une polarisation inverse V. Dans ce cas, si la diode est éclairée, sa courbe caractéristique inverse descend de I_{ph} comme l'indique la figure 11.21, *c* et dans la résistance de charge R_{ch} il apparaît un signal $\Delta V = I_{ph}R_{ch}$ qui est enregistré. A la différence d'une cellule photo-électrique, la photodiode possède une caractéristique courant-lumière linéaire et, en règle générale, présente des propriétés fréquentielles plus élevées. La rapidité de fonctionnement des photodiodes modernes atteint $\sim 10^{-9}$ s.

§ 11.9. Diodes luminescentes

Le passage d'un courant direct fait croître brusquement la concentration des porteurs minoritaires près des frontières de la jonction *p-n* qui devient nettement supérieure à la concentration d'équilibre. Aux porteurs minoritaires injectés s'ajoutent des porteurs majoritaires dont la concentration près des frontières de la jonction devient-elle aussi supérieure à la concentration d'équilibre. En diffusant vers les régions profondes du semiconducteur, les porteurs hors d'équilibre se recombinent en pénétrant en moyenne à une distance égale à la longueur de diffusion à partir de la couche de charge d'espace de la jonction *p-n*. Si une partie considérable des actes de recombinaison s'effectuent avec émission de la lumière, des conditions se réalisent pour la sortie de la lumière vers l'extérieur et donc la diode à semiconducteur peut être utilisée comme source de rayonnement lumineux. Une telle diode est appelée *diode luminescente* ou *diode émettrice de lumière*.

L'efficacité d'une diode luminescente se détermine avant tout par son *rendement quantique interne* η_{int} qui est égal au rapport du nombre de quanta émis lors de la recombinaison au nombre de porteurs minoritaires injectés. Si la recombinaison était uniquement radiative, on aurait $\eta_{int} = 1$. Or, une recombinaison radiative s'accompagne toujours d'une recombinaison non radiative. C'est pourquoi, dans le cas général, $\eta_{int} < 1$ et le rendement quantique se détermine par la relation suivante:

$$\eta_{int} = \frac{P_r}{P_r + P_{n.r}} = \frac{1}{1 + P_{n.r}/P_r}, \tag{11.52}$$

où P_r et $P_{n.r}$ sont respectivement les probabilités de recombinaisons radiative et non radiative rapportées à l'unité de temps.

La relation (11.52) montre que pour obtenir une efficacité interne maximale d'une diode luminescente, il faut augmenter autant que possible le rapport de la probabilité de recombinaison radiative à celle de recombinaison non radiative. La probabilité de recombinaison non radiative peut être diminuée en purifiant le semiconducteur des centres de recombinaison profonds. Cette opération est très difficile à réaliser parce que la section efficace de capture des porteurs par certains centres d'impureté, par exemple par le cuivre, est grande et le degré exigé de purification de telles impuretés est très élevé. Aussi, la qualité des diodes luminescentes dépend-elle pour une large part de la pureté des matériaux de départ et de la perfection technologique de la fabrication des diodes.

Un autre problème sérieux que pose la réalisation des diodes luminescentes est la sortie du rayonnement vers l'extérieur. Le fait est que l'indice de réfraction étant élevé ($\bar{n} = 3{,}5$), les semiconduc-

teurs ont un très petit angle de réflexion intérieure totale ($\Theta_{r.i.t}$):

$$\Theta_{r.i.t} = \arcsin(1/\bar{n}). \qquad (11.53)$$

C'est ainsi que dans le GaAs l'angle $\Theta_{r.i.t} \approx 17°$. C'est seulement ~ 2 % de rayonnement émis qui tombe sur la surface plane du semiconducteur sous des angles inférieures à $\Theta_{r.i.t}$ et peut sortir à l'extérieur en ne subissant qu'une réflexion partielle sur la surface frontière entre le semiconducteur et l'air. Le reste du rayonnement est réfléchi en totalité sur la surface à l'intérieur du semiconducteur. Si les coefficients d'absorption pour le rayonnement de recombinaison sont grands, la lumière réfléchie sur la surface sera absorbée dans le matériau de la diode. Si ces coefficients sont petits, on peut « sauver » une partie du rayonnement non sorti à l'extérieur en assurant sa réflexion sur les surfaces intérieures de la diode et en la dirigeant encore une fois sur la surface extérieure sous des angles inférieures à $\Theta_{r.i.t}$.

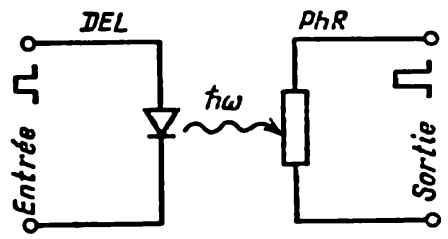

Fig. 11.22. Schéma d'une paire optronique:
DEL — diode émettrice de lumière; PhR — photorécepteur.

En radiotechnique, une cellule photoluminescente associée à un photorécepteur (*paire optronique*) peut servir de transformateur assurant un « découplage » électrique des circuits (fig. 11.22). La rapidité de fonctionnement des diodes luminescentes atteint $\sim 10^{-9}$ s. Les photorécepteurs peuvent assurer des constantes de temps de la photoréponse ayant elles aussi ces petites valeurs de sorte qu'un tel transformateur présente des caractéristiques de fréquence exceptionnelles. En utilisant une réaction galvanique ou optique et la non-linéarité des caractéristiques des photorécepteurs et des diodes luminescentes, on peut réaliser à l'aide des optrons des *éléments logiques* variés.

La puissance moyenne émise en continu par une diode luminescente n'est pas élevée — de l'ordre de quelques milliwatts. Son augmentation par l'élévation du courant direct est impossible à réaliser à cause d'un échauffement anormal de la jonction *p-n* qui provoque une brusque réduction du rendement quantique interne. Néanmoins les puissances même de si petites valeurs permettent d'utiliser les diodes luminescentes pour la réalisation des *liaisons optiques* directionnelles dans les limites de la visibilité directe. Comme les dimensions de la région luminescente de la diode sont petites, le rayonnement peut être facilement réalisé sous forme de faisceau faiblement divergent. C'est aussi l'utilisation des diodes luminescentes comme source de rayonnement dans des lignes de télécommunications à fibre optique qui semble vouée à un avenir prometteur.

Les diodes luminescentes sont aussi employées comme *indicateurs lumineux* de faible encombrement. Dans ce cas, un des paramètres principaux de la diode luminescente est sa brillance telle qu'elle est perçue par l'œil de l'homme. En utilisant pour la réalisation d'une diode luminescente des semiconducteurs à différente largeur de la bande interdite, on peut obtenir différentes couleurs de la luminescence de la diode.

§ 11.10. Lasers à semiconducteurs

Emission stimulée (induite). Dans les *lasers* qui sont des appareils servant à la génération d'un rayonnement lumineux monochromatique étroitement guidé on utilise le phénomène d'*émission stimulée de la lumière*. Ce phénomène consiste en passage d'un électron du

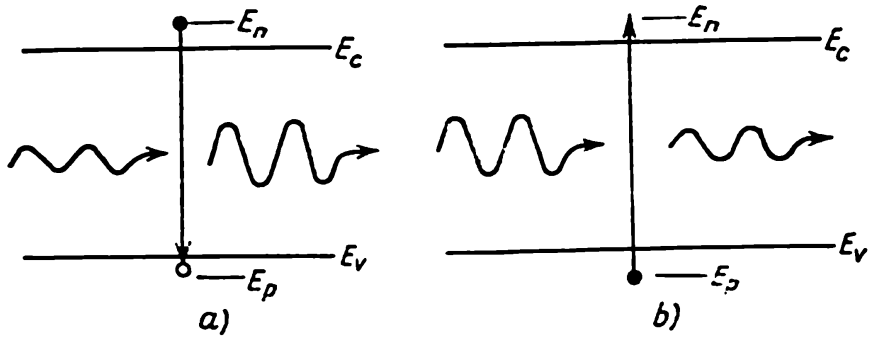

Fig. 11.23. Schéma de l'émission stimulée (*a*) et d'absorption (*b*) de la lumière.

niveau énergétique supérieur E_n (fig. 11.23, *a*) à un niveau inférieur libre E_p sous l'action d'une onde lumineuse de fréquence

$$\nu = (E_n - E_p)/h. \tag{11.54}$$

L'énergie $E_n - E_p$ qui s'en trouve libérée est transférée à l'onde lumineuse qui a provoquée la transition, en augmentant son amplitude. Il se produit une *amplification de la lumière*.

L'émission stimulée est l'inverse du processus d'absorption de la lumière quand un électron passe du niveau inférieur E_p au niveau supérieur libre E_n en absorbant l'énergie de l'onde lumineuse (fig. 11.23, *b*).

Si, dans un milieu, les électrons occupent les états tant supérieurs qu'inférieurs, l'amplification de la lumière l'emportera sur son absorption si la population des niveaux supérieurs est plus élevée que celle des niveaux inférieurs (inversion de population). Dans le cas contraire, il y aura absorption de la lumière. Dans les conditions d'équilibre thermique, les populations des états d'énergie inférieurs sont toujours plus élevées que celles des états d'énergie supérieurs de sorte que la lumière est toujours absorbée par de tels systèmes. Pour obtenir une inversion de population, il faut faire sortir le système de

l'équilibre thermique. Le milieu dans lequel on a réalisé l'inversion de population des niveaux est dit *actif*.

Dans les lasers, l'onde lumineuse effectue de multiples parcours à travers le milieu actif en se réfléchissant sur les miroirs (fig. 11.24). Si l'énergie perdue par l'onde lumineuse par suite de sa réflexion sur un miroir est inférieure à l'énergie qu'elle acquiert en passant par le milieu actif, l'intensité de l'onde augmentera en permanence. Il se produit une augmentation en avalanche de l'intensité de l'onde lumineuse : on dit que le laser entre en oscillation. Vu qu'à chaque réflexion une partie de la lumière passe à travers le miroir, l'intensité de la lumière sortant du laser à l'extérieur augmente en même temps

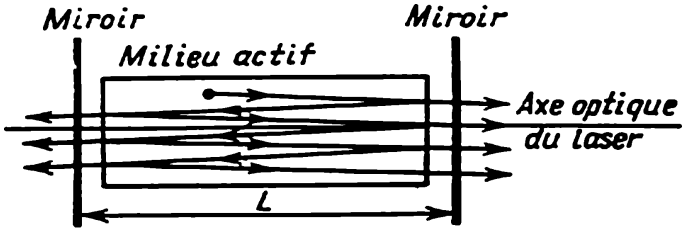

Fig. 11.24. Schéma d'un parcours multiple de la lumière dans le milieu actif d'un laser.

que l'intensité de l'onde lumineuse dans le milieu actif. Pourtant, l'augmentation de l'amplitude de l'onde lumineuse provoque une augmentation de la cadence des transitions stimulées des électrons du niveau supérieur au niveau inférieur et donc conduit à une diminution de l'inversion de population et à une diminution de l'amplification de la lumière par un parcours du milieu actif de sorte que finalement il s'établit une amplitude stationnaire de l'onde lumineuse pour laquelle l'amplification de la lumière pendant un seul parcours du milieu actif est égale aux pertes.

Lasers à semiconducteurs. Les lasers à semiconducteurs utilisant le phénomène d'émission stimulée qui provient de la recombinaison radiative des électrons et des trous ont une grande importance pour la résolution de nombreux problèmes qui se posent devant la radioélectronique et la technique de calcul électronique. Les plus employés en pratique sont des *lasers à injection* à l'arséniure de gallium dans lesquels l'inversion de population est obtenue par injection de porteurs minoritaires à travers une jonction *p-n* dans des régions dégénérées du semiconducteur. On utilise aussi l'arséniure d'indium, le phosphure d'indium, l'antimoniure d'indium ainsi que certaines solutions solides, par exemple InGaAsP, AlGaAsP, etc.

La figure 11.25, *a* montre une jonction *p-n* en équilibre entre deux régions dégénérées d'un semiconducteur. Le niveau de Fermi dans la région *p* (μ) est situé plus bas que le sommet de la bande de valence E_v. Une telle position du niveau de Fermi témoigne que les états au

voisinage du sommet de la bande de valence de la région p sont, avec une probabilité voisine de 1, libres (occupés par des trous). Si l'on applique à une telle jonction p-n une polarisation directe V qui réduit brusquement la barrière de potentiel, il apparaîtra dans cette jonction une région A à inversion de population des bandes : au-dessus des niveaux pratiquement libres de la bande de valence se situent les niveaux occupés de la bande de conduction (fig. 11.25, b). Dans ces conditions, les quanta, nés de façon spontanée par suite de la recombinaison des paires électron-trou, provoquent une émission stimulée d'un rayonnement.

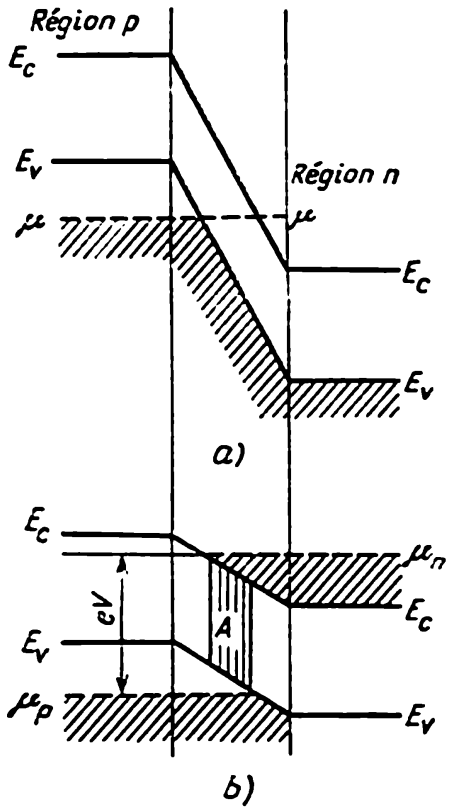

Fig. 11.25. Diagramme énergétique d'une diode laser à l'état d'équilibre (a) et en fonctionnement (b) (μ_n et μ_p — quasi-niveaux de Fermi pour les électrons et les trous).

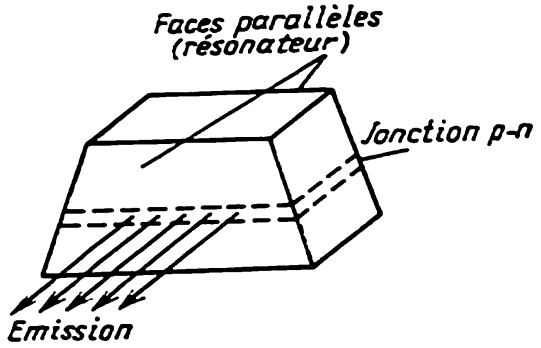

Fig. 11.26. Schéma de constitution d'un laser à injection à semiconducteur.

C'est sur ce principe qu'est fondé le fonctionnement des lasers à semiconducteurs dont le schéma de constitution est représenté sur la figure 11.26. Un cristal comportant une jonction p-n a la forme d'un parallélépipède ou d'une pyramide irrégulière : deux faces opposées sont rendues rigoureusement parallèles l'une à l'autre et perpendiculaires au plan de la jonction p-n ; elles jouent le rôle d'un résonateur optique obligeant l'émission stimulée qui a pris naissance dans le plan de la jonction de le traverser beaucoup de fois. Les deux autres faces peuvent être orientées sous un angle par rapport à la base et restent grossièrement usinées de sorte qu'elles ne peuvent pas jouer le rôle d'un résonateur optique. Le rayonnement cohérent sort par une des faces du résonateur optique.

Le coefficient de réflexion de la lumière sur les faces du cristal est de 0,3 à 0,35. En outre, l'onde lumineuse qui se propage le long

de la jonction *p-n* passe non seulement par la région active mais également par des régions passives de la diode. C'est pourquoi, pour rendre possible l'oscillation du laser, il est nécessaire de réaliser une inversion de population des bandes telle que l'amplification de la lumière dans le milieu actif dépasse toutes les pertes liées au passage à travers la diode et à la faible réflexion sur les faces-miroirs.

Le courant I_{seuil} pour lequel cette condition est satisfaite et le laser entre en oscillation est appelé *courant de seuil*. Tant que le courant n'a pas atteint sa valeur de seuil, le laser fonctionne comme une diode luminescente ordinaire et effectue une émission spontanée dont la densité est la même dans toutes les directions (dans un angle solide de 4π stéradians). Les rayons qui ne sont pas tombés sur les faces réflectrices du cristal y sont totalement absorbés. En outre, les rayons tombant sur ces faces sous un angle $\Theta > 17°$ subissent une réflexion interne totale et finalement sont eux aussi absorbés dans le cristal. Pour toutes ces raisons il ne sort de la diode luminescente que près de 2 % de tout le rayonnement qui y est produit par suite d'une recombinaison radiative.

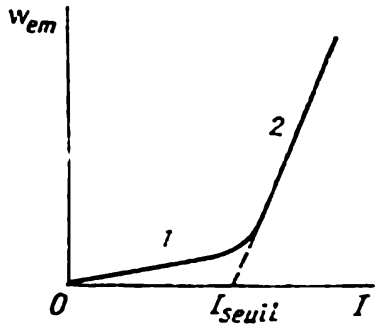

Fig. 11.27. Variation de la puissance émise par une diode laser en fonction du courant:
1 — émission spontanée; 2 — émission stimulée.

Après le passage au régime d'oscillation, tout le rayonnement est concentré pratiquement dans le plan de la jonction *p-n* en se propageant perpendiculairement aux faces réflectrices. En outre, pour $I > I_{seuil}$ la probabilité de transitions optiques stimulées augmente, ce qui fait croître le rapport des probabilités de la recombinaison radiative et de la recombinaison non radiative. Tout cela conduit à une croissance brusque de la puissance émise $W_{ém}$ et à une cassure de la courbe de variation de $W_{ém}$ en fonction du courant I pour $I = I_{seuil}$ (fig. 11.27).

Un inconvénient sérieux des lasers à semiconducteurs est que leurs paramètres dépendent fortement de la température. Lorsque la température augmente par suite de l'échauffement de la diode parcourue par un courant direct important, la largeur de la bande interdite varie, ce qui conduit à une modification de la composition spectrale du rayonnement et au déplacement de son maximum vers des longueurs d'onde plus grandes. Mais ce qui compte principalement c'est la croissance du courant de seuil I_{seuil} avec la température parce qu'à un courant d'injection inchangé et donc à une concentration invariable des porteurs injectés près de la jonction *p-n* leur répartition en fonction de l'énergie devient plus floue — il se produit une aug-

mentation de l'intervalle d'énergie (de l'ordre de kT) dans les limites duquel les porteurs de charges libres sont répartis dans les bandes d'énergie. Comme le coefficient d'amplification de la lumière dépend du degré de remplissage par les électrons et par les trous des états dans la bande de conduction et la bande de valence respectivement, les coefficients d'amplifications diminuent, lorsque la température augmente pour un même niveau d'injection. Cela signifie que pour obtenir une valeur de seuil du coefficient d'amplification aux températures élevées, il est nécessaire d'assurer un courant de seuil I_{seuil} bien intense. Aussi le problème de l'évacuation de la chaleur dégagée dans la jonction *p-n* a-t-il pour les lasers à semiconducteur une importance primordiale.

Applications de l'émission cohérente. Un haut degré de monochromatisme et une faible divergence de l'émission lumineuse cohérente déterminent les domaines de son emploi pratique. L'émission de haute cohérence temporelle peut être utilisée pour la transmission de l'information à des fréquences optiques, pour la résolution des problèmes liés à l'interférence optique (mesure des distances, des vitesses linéaire et angulaire, des déformations de surfaces, etc.) et comme standard de fréquence. Une haute directivité de l'émission spatialement cohérente détermine toute une série d'avantages sur une émission non cohérente: une petitesse de pertes d'énergie, liées à la divergence du faisceau, une haute résolution angulaire permettant de diriger le rayon exactement sur un petit objet et de réduire considérablement les parasites, la possibilité de filtrage spatial des signaux. Cela signifie que l'émission lumineuse hautement directionnelle peut être utilisée de façon efficace pour la transmission de l'information à de grandes distances, pour la détection optique des objets éloignés (surtout pour distinguer un objet parmi d'autres cibles), pour la mesure des angles et des distances suivant un principe sur lequel est fondé le fonctionnement des géodimètres utilisés dans les systèmes de guidage par faisceau, pour la transmission de l'énergie lumineuse à de grandes distances, pour la concentration de l'énergie dans un faisceau lumineux (coupe optique, soudage optique), etc.

Les lasers à semiconducteurs à couplage optique peuvent être utilisés pour la réalisation des éléments logiques optiques à haute rapidité de fonctionnement. Actuellement on a réalisé pratiquement des éléments logiques à lasers à semiconducteurs dont la rapidité de fonctionnement est de 10^{-10} s. Sur la base de ces éléments on peut fabriquer des dispositifs pour les calculateurs électroniques possédant une très haute rapidité de fonctionnement, un découplage maximal entre l'entrée et la sortie ainsi qu'une haute immunité contre les brouillages.

CHAPITRE 12

PRINCIPES PHYSIQUES DE FONCTIONNEMENT DES TRANSISTORS

L'élément principal et le plus important de l'électronique des semiconducteurs est le *transistor* qui est un appareil à état solide destiné à l'amplification de signaux électriques, à la conversion de l'énergie d'une source d'alimentation en énergie des oscillations électromagnétiques (production d'oscillations), à une commutation rapide et à la réalisation de nombreuses autres fonctions importantes. De nos jours, les transistors constituent les composants les plus importants des calculatrices à grande vitesse, pratiquement de tous les matériels de télécommunication et des appareils électroniques domestiques. Des transistors bipolaires ont été inventés en 1947. En 1948 on a réalisé un transistor à pointes. En 1949 Bardeen a publié un ouvrage consacré aux transistors à jonctions. En 1952 on a proposé un transistor à effet de champ à jonction *p-n* et puis un transistor à effet de champ à grille isolée.
Dans ce chapitre vont être étudiés les principes physiques de fonctionnement de divers types de transistors bipolaires et à effet de champ.

§ 12.1. Principe de fonctionnement d'un transistor bipolaire

Un transistor bipolaire est constitué par deux jonctions *p-n* écartées d'une petite distance. Un tel transistor peut être, par sa structure, soit du type *p-n-p*, soit du type *n-p-n*. Considérons par exemple un transistor du type *n-p-n* (fig. 12.1, *a*). Son diagramme énergétique pour le cas où il est à l'état d'équilibre, c'est-à-dire non soumis à aucune tension extérieure, est représenté sur la figure 12.1, *b*. La région centrale du transistor est appelée *base* (*B*) et les deux régions latérales sont appelées *émetteur* (*E*) et *collecteur* (*C*).

Examinons le régime de fonctionnement d'un transistor dans un *montage* dit *à base commune* (fig. 12.2, *a*). La jonction émettrice est polarisée en sens direct et la jonction collectrice en sens inverse. Les électrons de la région émettrice sont injectés à travers la barrière de potentiel abaissée de la jonction émettrice (flèches *I'* des figures 12.2, *a* et *b*) dans la région de base du transistor. Pour compenser la charge des électrons injectés, un nombre nécessaire de trous arrive vers ces électrons par le contact ohmique de base. De même que dans une diode ordinaire (v. § 11.2), la polarisation extérieure V_ℓ chute pratiquement en totalité dans la région de charge d'espace de la jonction émetteur-base et la tension V_c le fait dans la jonction base-collecteur et dans la résistance extérieure R_{ch}. Ainsi, le champ électri-

que dans la base n'existe pas et les électrons, injectés dans la base à travers la jonction émettrice, diffusent vers la jonction collectrice. Leur concentration près de la jonction émettrice est exp $[eV_é/(kT)]$ fois plus grande que la concentration d'équilibre tandis que près de la frontière da la jonction *p-n* collectrice, polarisée en sens inverse, elle est pratiquement nulle. Si l'épaisseur w de la région de base est très inférieure à la longueur de diffusion des électrons dans la base, les électrons atteignent la jonction collectrice pratiquement sans se recombiner dans la base.

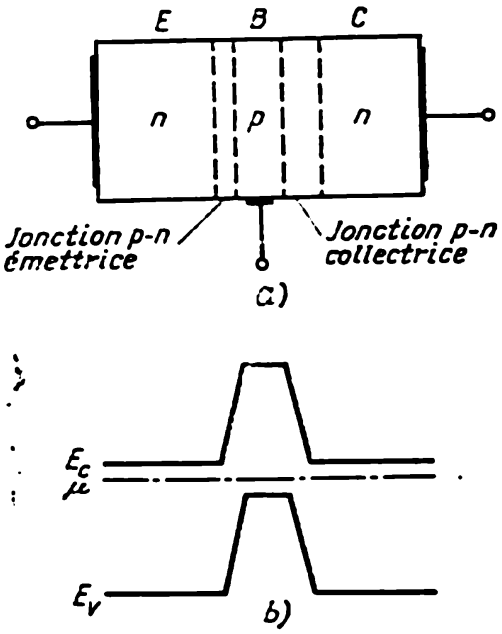

Fig. 12.1. Structure d'un transistor *n-p-n* (*a*) et son diagramme énergétique (*b*).

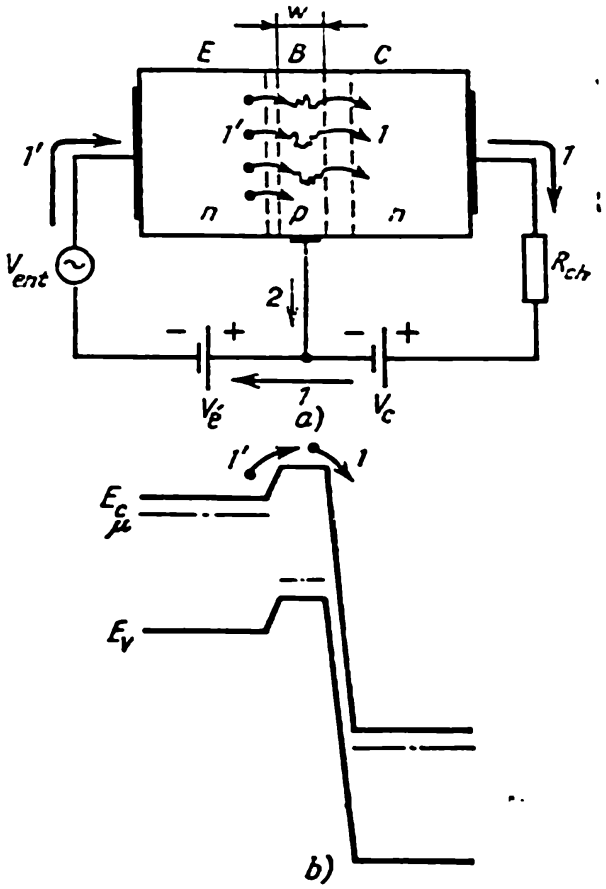

Fig. 12.2. Fonctionnement d'un transistor monté en base commune (*a*) et son diagramme énergétique en état de régime (*b*).

Quand les électrons pénètrent, par suite de l'agitation thermique désordonnée, dans la région de charge d'espace de la jonction collectrice, ils sont entraînés par le champ de cette jonction qui les fait passer dans la région collectrice du transistor (flèches *1* de la figure 12.2). Ainsi, le courant électronique de la jonction émettrice $I_{né}$ du transistor passe non par la sortie de base mais à travers le collecteur et ensuite par la résistance de charge R_{ch}. Quant au courant passant par la sortie de base, il est beaucoup moins intense que le courant d'émetteur, parce que le courant de base ne sert qu'à compenser la perte de trous qui se recombinent avec une petite partie des électrons diffusant à travers la base. L'utilisation d'un transistor

dans le montage à base commune permet d'obtenir une amplification de la tension dans le circuit de sortie du transistor (dans la résistance de charge R_{ch}) par rapport à la tension V_{ent} agissant dans son circuit d'entrée. En effet, une tension variable de faible valeur V_{ent} appliquée au circuit de la jonction émettrice polarisée en sens direct fait varier considérablement son courant (le courant de la jonction *p-n* polarisée en sens direct varie de *e* fois lorsque la polarisation varie de $kT/e \approx 25$ mV). En parcourant la résistance de charge R_{ch}, ce courant y développe, si la valeur ohmique de la résistance est suffisamment forte, une tension beaucoup plus grande que V_{ent}. Par voie de conséquence, la puissance de courant alternatif dégagée dans la charge augmente elle aussi par rapport à la puissance consommée à l'entrée du transistor. Dans ces conditions, pour que la jonction collectrice reste polarisée en sens inverse malgré le passage du courant dans la résistance de charge, il faut que la tension V_c de la source de polarisation soit suffisamment forte.

§ 12.2. Paramètres et caractéristiques de sortie d'un transistor

Paramètres d'un transistor. La caractéristique la plus importante d'un transistor est son *gain en courant* α défini par le rapport de la variation du courant de collecteur à la variation du courant d'émetteur qui l'a provoquée, à tension constante appliquée à la jonction *p-n* collectrice (c'est-à-dire à $R_{ch} = 0$ sur la figure 12.2, *a*):

$$\alpha = (\partial I_c/\partial I_\acute{e})_{V_c}. \tag{12.1}$$

Le coefficient α est appelé *paramètre externe* d'un transistor. Il se détermine par trois *paramètres internes*: *l'efficacité de l'émetteur* γ, *le coefficient de transfert* β, *l'efficacité du collecteur* α^*.

Le courant d'émetteur vaut la somme d'une composante électronique ($I_{n\acute{e}}$) et d'une composante de trous ($I_{p\acute{e}}$). Dans un transistor du type *n-p-n* l'effet amplificateur n'est obtenu que grâce à la composante électronique du courant d'émetteur. L'efficacité γ de l'émetteur est égale au rapport de la variation de la composante électronique du courant d'émetteur à la variation du courant total d'émetteur $I_\acute{e}$:

$$\gamma = (\partial I_{n\acute{e}}/\partial I_\acute{e})_{V_c}. \tag{12.2}$$

Suivant (11.27) et (11.28) on a

$$I_{p\acute{e}} = S_\acute{e} e \frac{p_{n0}}{\tau_p} L_p \{\exp[eV_\acute{e}/(kT)] - 1\} \approx$$

$$\approx S_\acute{e}\, e \frac{p_{n0}}{\tau_p} L_p \exp[eV_\acute{e}/(kT)], \tag{12.3}$$

où $S_\acute{e}$ est la surface de la jonction émettrice.

En faisant usage de la relation $L = \sqrt{D\tau}$, on peut mettre l'expression (12.3) aussi sous la forme suivante:

$$I_{p\acute{e}} \approx S_{\acute{e}} e D_p \frac{p_{n0}}{L_p} \exp\left[eV_{\acute{e}}/(kT)\right]. \tag{12.4}$$

La composante électronique du courant d'émetteur se détermine par une formule autre que (12.3). Le fait est que si les trous injectés de la base dans l'émetteur se répartissent à partir de la jonction à une profondeur L_p, les électrons injectés de l'émetteur dans la base ne s'y propagent qu'à une distance $w < L_n$ (où w est l'épaisseur de la région de base). Vu que les électrons se déplacent dans la base seulement par diffusion, la valeur de la composante électronique du courant d'émetteur se détermine par le gradient de concentration des électrons

$$I_{n\acute{e}} = S_{\acute{e}} e D_n \left| \frac{dn}{dx} \right|, \tag{12.5}$$

où l'axe des x est dirigé de l'émetteur vers le collecteur.

Proposons-nous de calculer le gradient de concentration des électrons dans la base. Près de la jonction émettrice leur concentration a pour valeur

$$n(0) = n_{p0} \exp \frac{eV_{\acute{e}}}{kT}, \tag{12.6}$$

et près de la jonction collectrice $n(w) = 0$. Par conséquent, il vient

$$\left| \frac{dn}{dx} \right| \approx \frac{n(0) - n(w)}{w} = \frac{n_{p0}}{w} \exp \frac{eV_{\acute{e}}}{kT}. \tag{12.7}$$

En portant (12.7) dans (12.5), on obtient

$$I_{n\acute{e}} = S_{\acute{e}} e D_n \frac{n_{p0}}{w} \exp \frac{eV_{\acute{e}}}{kT}. \tag{12.8}$$

En tenant compte que $I_{\acute{e}} = I_{n\acute{e}} + I_{p\acute{e}}$ et en combinant (12.2), (12.4) et (12.8), on obtient pour l'efficacité de l'émetteur l'expression suivante:

$$\gamma = \left(1 + \frac{p_{n0}}{n_{p0}} \frac{D_p}{D_n} \frac{w}{L_p}\right)^{-1}. \tag{12.9}$$

Si l'on exprime les coefficients de diffusion en fonction des mobilités à l'aide de la relation d'Einstein (8.42) et les concentrations des porteurs minoritaires en fonction des concentrations des porteurs majoritaires au moyen de la loi de l'action de masse (5.36), on peut récrire l'expression (12.9) sous une forme plus commode pour les ap-

plications pratiques:

$$\gamma = \left(1 + \frac{\sigma_p}{\sigma_n}\frac{w}{L_p}\right)^{-1} \approx 1 - \frac{\sigma_p}{\sigma_n}\frac{w}{L_p}. \qquad (12.10)$$

Pour rendre l'efficacité de l'émetteur plus proche de l'unité il faut, comme le montre l'expression (12.10), doper l'émetteur beaucoup plus fortement que la base ($\sigma_n \gg \sigma_p$) et rendre la base suffisamment mince ($w \ll L_p$). Cette dernière exigence est aussi nécessaire pour augmenter le coefficient de transfert β égal à la partie des électrons qui sont injectés de l'émetteur et ont atteint la jonction collectrice:

$$\beta = (\partial I_{nc}/\partial I_{n\acute{e}})_{V_c}. \qquad (12.11)$$

Ce coefficient est inférieur à l'unité parce que certains des électrons cheminant vers le collecteur se recombinent. Plus étroite est la base par rapport à la longueur de diffusion des électrons, plus voisin de l'unité est le coefficient β. Les calculs montrent que pour une base étroite ($w \ll L_n$)

$$\beta \approx 1 - \frac{1}{2}\left(\frac{w}{L_n}\right)^2. \qquad (12.12)$$

L'efficacité du collecteur définie par

$$\alpha^* = (\partial I_c/\partial I_{nc})_{V_c}, \qquad (12.13)$$

peut dépasser l'unité si dans la jonction *p-n* collectrice il se produit une ionisation par choc. Quant au régime habituel, quand $V_c \ll V_{cl}$, la valeur de α est pratiquement égale à l'unité.

Des expressions (12.1), (12.2), (12.11) et (12.13) il résulte que

$$\alpha = \gamma\beta\alpha^*. \qquad (12.14)$$

En y portant les valeurs de γ et β données par (12.10) et (12.12), en posant $\alpha^* = 1$ et en rejetant les termes du deuxième ordre de petitesse, on obtient

$$\alpha \approx 1 - \frac{\sigma_p}{\sigma_n}\frac{w}{L_p} - \frac{1}{2}\left(\frac{w}{L_n}\right)^2. \qquad (12.15)$$

Comme le montre la relation (12.15), dans le montage à base commune le courant n'est pas amplifié ($\alpha < 1$).

Les caractéristiques de sortie d'un transistor (fig. 12.3) traduisent la variation du courant de collecteur en fonction de la tension collecteur-base. Pour leur construction on considère d'habitude que la polarisation inverse et le courant inverse de la jonction collectrice sont positifs. Ces caractéristiques présentent une certaine pente positive due à ce que le gain α augmente avec V_c. Le fait est que l'augmentation de la polarisation inverse provoque un élargissement de la

région de charge d'espace de la jonction *p-n* collectrice et un *rétrécissement de la base w* du transistor. C'est ce qui conduit, suivant (12.15), à une croissance de α.

Le coefficient α dépend aussi du courant d'émetteur. Cette dépendance est représentée schématiquement sur la figure 12.4. De faibles courants d'émetteur correspondent à de petites tensions sur la jonction *p-n* émettrice quand une partie du courant direct est due non pas à l'injection des porteurs à travers la jonction mais à leur recombinaison dans la couche de charge d'espace de la jonction *p-n*. Vu

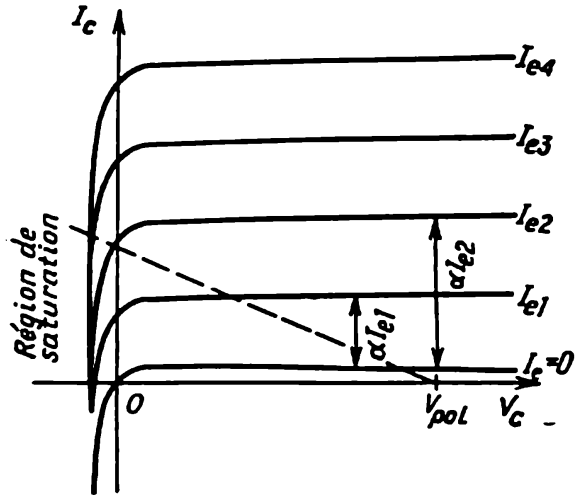

Fig. 12.3. Caractéristiques de sortie d'un transistor monté en base commune.

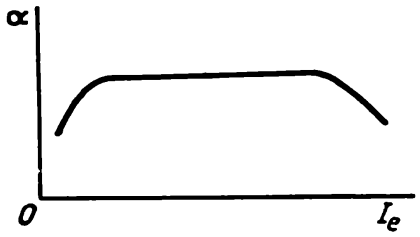

Fig. 12.4. Variation du gain en courant α d'un transistor en fonction du courant d'émetteur.

que l'effet transistor n'est assuré que par des porteurs injectés dans la base, la recombinaison des porteurs dans la jonction *p-n* conduit à une réduction de l'efficacité γ de l'émetteur et du gain en courant α.

Vers les porteurs minoritaires injectés arrivent, par le circuit de base, des porteurs majoritaires pour compenser la charge d'espace injectée. Dans le cas des courants d'émetteur intenses, le nombre de porteurs minoritaires (d'électrons dans un transistor *n-p-n*) injectés dans la base est si grand que la concentration des porteurs majoritaires près de l'émetteur augmente notablement. Il en résulte une croissance du courant d'injection des porteurs majoritaires de la base dans l'émetteur, c'est-à-dire une diminution de l'efficacité γ de l'émetteur.

Lorsque le circuit de collecteur comporte une résistance de charge R_{ch}, la tension agissant sur la jonction collectrice varie avec le courant de collecteur. Ces variations peuvent être déterminées à l'aide de la droite de charge (ligne en traits interrompus de la figure 12.3). Si le nombre de porteurs injectés dans la base est si grand que

la chute de tension $I_c R_{ch}$ développée dans la charge dépasse la tension de la source de polarisation V_{pol}, la jonction *p-n* collectrice sera polarisée en sens direct. On dit dans ce cas que le transistor fonctionne dans la région de saturation. Bien qu'en ce régime, dans la jonction *p-n* collectrice il existe toujours un champ et donc les porteurs minoritaires injectés de l'émetteur dans la base sont poussés par ce champ vers le collecteur, un flux opposé de porteurs cheminemaintenant à travers la barrière de faible hauteur de la jonction *p-n*.

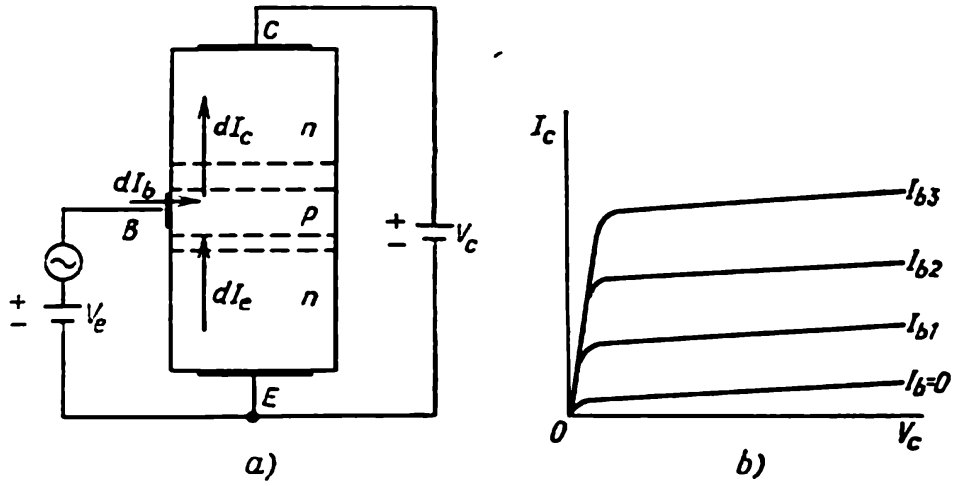

Fig. 12.5. Schéma de branchement (*a*) et caractéristiques de sortie (*b*) d'un transistor monté en émetteur commun.

Comme il est visible sur la figure 12.3, l'augmentation du courant d'émetteur laisse pratiquement inchangée la tension de collecteur.

Montage à émetteur commun. Bien souvent le transistor est monté comme l'indique la figure 12.5, *a*. Dans ce montage, il fonctionne en amplificateur de courant. Le courant d'entrée est ici le courant de base I_b et non pas celui d'émetteur. Les variations des courants de base, d'émetteur et de collecteur sont liées entre elles par la relation

$$dI_c = dI_é - dI_b. \tag{12.16}$$

D'un autre côté,

$$dI_c = \alpha dI_é. \tag{12.17}$$

En combinant (12.16) avec (12.17), on obtient

$$dI_c = \frac{\alpha}{1-\alpha} dI_b = B_0 dI_b, \tag{12.18}$$

où B_0 désigne le *gain en courant dans le montage à émetteur commun*. Comme dans le transistor $\alpha \approx 1$, ce gain $B_0 \gg 1$. Par exemple, dans un transistor de bonne qualité $\alpha = 0{,}995$. Alors, suivant (12.18) $B_0 = 200$.

Expliquons la nature physique de l'amplification du courant dans le montage à émetteur commun. Supposons qu'une certaine quantité ΔN d'électrons est injectée dans la base à travers une jonction émettrice polarisée en sens direct. Pour compenser leur charge, une quantité $\Delta P = \Delta N$ de trous est arrivée dans la base. Comme les électrons injectés par l'émetteur participent à l'agitation thermique désordonnée, ils atteignent la jonction *p-n* collectrice et sont attirés par son champ. Leur place est occupée par de nouveaux électrons en provenant de l'émetteur et ainsi de suite. Ce processus continue tant que les trous ΔP introduits dans la base ne se recombinent pas avec les électrons qui diffusent à travers la base depuis l'émetteur. Chaque trou passant par le contact de base fait traverser les jonctions émettrice et collectrice à un nombre d'électrons autant de fois plus grand que la durée de vie des électrons dans la base est plus grande que la durée de leur dérive à travers la base.

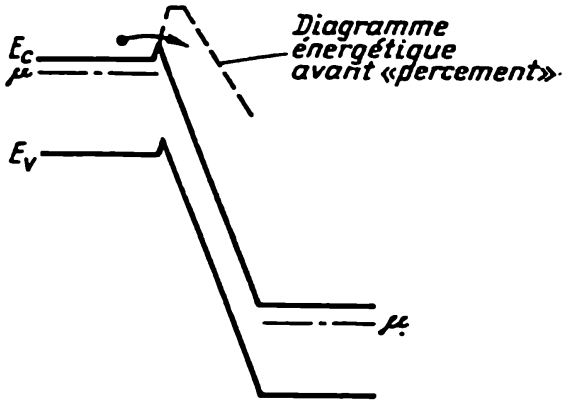

Fig. 12.6. Diagramme énergétique d'un transistor *n-p-n* à base « percée ».

Les caractéristiques de sortie d'un transistor monté en émetteur commun sont représentées sur la figure 12.5, *b*. Leur pente est nettement plus grande que dans le montage à base commune parce qu'une petite augmentation de α, lorsque V_c augmente par suite du rétrécissement de la base, a pour effet, suivant (12.18), une croissance beaucoup plus grande de B_0.

Un *claquage* de la jonction *p-n* collectrice peut se produire de la même façon que dans une diode ordinaire. En outre, une croissance brutale du courant de collecteur peut être provoquée par le « percement de la base » du transistor quand les couches de la charge d'espace des jonctions *p-n* collectrice et émettrice se rejoignent. Dans ce cas, si la tension de collecteur s'accroît encore, la barrière de potentiel de la jonction émettrice s'abaisse comme l'indique la figure 12.6 et donc le courant venant de l'émetteur dans le collecteur s'accroît.

§ 12.3. Processus transitoires dans les transistors. Transistors à dérive

Examinons les processus transitoires dans les transistors sur l'exemple d'un transistor *n-p-n* monté en base commune (fig. 12.7, *a*). Faisons passer par la jonction émettrice une impulsion de courant de forme rectangulaire (fig. 12.7, *b*). L'impulsion du courant de collecteur n'atteint sa valeur stationnaire, déterminée par les caractéris-

tiques de sortie du transistor, qu'au bout d'un certain temps qui est celui du processus transitoire (fig. 12.7, *c*). De même, après la fin de l'impulsion du courant d'émetteur le courant de collecteur ne diminue pas instantanément mais progressivement. La durée des processus transitoires se détermine par le temps de recharge des capacités de barrière des jonctions *p-n* émettrice et collectrice ainsi que par le temps de diffusion des porteurs injectés à travers la base.

L'augmentation du courant d'émetteur doit s'accompagner de l'augmentation de la polarisation directe de la jonction *p-n* émettrice et donc de la diminution de la largeur de la région de sa charge

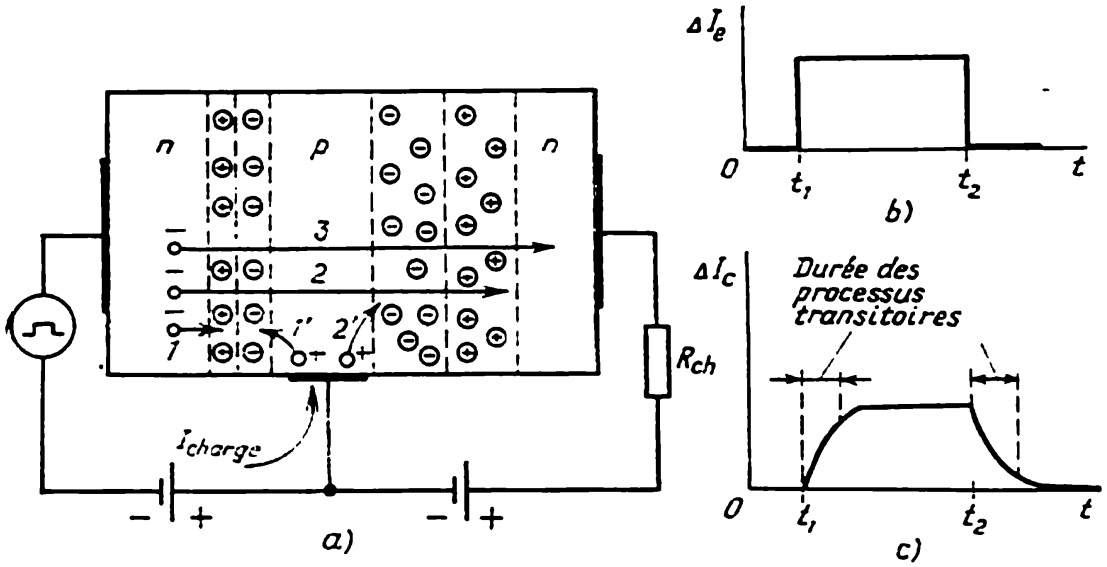

Fig. 12.7. Représentation schématique des processus transitoires dans un transistor (*a*); impulsion du courant d'émetteur ΔI_e (*b*); impulsion du courant de collecteur ΔI_c (*c*).

d'espace. Comme il a été montré au § 11.2, le rétrécissement de la région de charge d'espace de la jonction *p-n* se produit par suite du déplacement vers cette jonction des électrons de la région *n* et des trous de la région *p* (flèches *1* et *1'* de la figure 12.7, *a*). Ainsi, au début de l'impulsion, le courant d'émetteur est dépensé non pour l'injection des électrons à travers la jonction *p-n*, ce qui doit provoquer l'apparition d'un courant de collecteur, mais pour la compensation de la charge d'espace des impuretés près des frontières de la jonction, c'est-à-dire pour la charge de la capacité de barrière de la jonction *p-n* émettrice. Le temps de charge de cette capacité est égal à $R_{série}C_{b.e}$, où $C_{b.e}$ est la capacité de barrière de la jonction *p-n* émettrice et $R_{série}$ la résistance des régions d'émetteur et de base du transistor, de leurs contacts et la résistance extérieure du circuit d'émetteur. La résistance série propre d'un transistor est souvent déterminée par la résistance de sa région de base parce que pour arriver vers la jonction *p-n* émettrice pour charger sa capacité de barrié-

re les porteurs de courant venant du contact de base ont à traverser une région de base étroite (courant I_{ch} de la figure 12.7, *a*).

Au fur et à mesure que la hauteur de la barrière de potentiel de la jonction *p-n* émettrice diminue, les électrons de la région émettrice commencent à être injectés dans la base et diffuser vers le collecteur (flèche *3*). Comme le processus de diffusion est conditionné par une agitation thermique désordonnée des porteurs, certains parmi eux arrivent à atteindre le collecteur très vite tandis que d'autres mettent plus de temps pour cela. C'est pourquoi, si le courant d'émetteur était devenu d'emblée un courant d'injection, le courant de collecteur croîtrait progressivement pendant un temps égal au temps caractéristique de diffusion de la masse moyenne des porteurs à travers la région de base. Evaluons le temps de diffusion en partant de ce que pendant la durée de vie τ les porteurs minoritaires diffusent en moyenne à une distance égale à la longueur de diffusion L qui est liée à τ par la relation

$$L = \sqrt{D\tau}.$$

Il en résulte que les électrons diffusent à une profondeur w durant le temps

$$t_{dif} = w^2/D_n. \tag{12.19}$$

Comme $w \ll L_n$, on a $t_{dif} \ll \tau$.

Lorsque les électrons atteignent la jonction collectrice, ils subissent l'action de son champ qui les pousse dans la région collectrice. Cela ne signifie pourtant pas que le circuit de collecteur sera tout de suite parcouru par un courant correspondant ΔI_c. Le fait est que l'augmentation du courant de collecteur de ΔI_c doit provoquer une augmentation de $R_{ch}\Delta I_c$ de la chute de tension aux bornes de la charge. De cette même quantité doit diminuer la polarisation inverse de la jonction *p-n* collectrice, c'est-à-dire que la couche de sa charge d'espace doit se rétrécir. Ce rétrécissement se produit du fait que les électrons transférés à travers la jonction collectrice ne s'en vont pas dans le circuit extérieur mais restent près de la jonction *p-n* en réduisant sa charge d'espace du côté de la région *n* de collecteur (flèche *2* de la figure 12.7, *a*). Du côté de la base, la charge d'espace diminue à cause des trous arrivant à travers le contact de base depuis le circuit d'émetteur (flèche *2'* de la figure 12.7, *a*). La variation de la charge sur la jonction collectrice doit finalement prendre la valeur suivante:

$$\Delta Q = C_{b.c}\Delta V_{ch} = C_{b.c}R_{ch}\Delta I_c. \tag{12.20}$$

Comme cette variation est due au courant ΔI_c, elle se produira pendant le temps

$$t_{charge} = \Delta Q/\Delta I_c = C_{b.c}R_{ch}. \tag{12.21}$$

Ici, $C_{b.c}$ est la capacité de barrière de la jonction *p-n* collectrice.

Après la fin de l'impulsion du courant d'émetteur le courant dans le circuit de collecteur est entretenu grâce à la décharge, à travers ce circuit, des capacités de barrière des jonctions émettrice et collectrice ainsi que grâce à l'arrivée vers le collecteur des porteurs injectés pendant la durée d'action de l'impulsion.

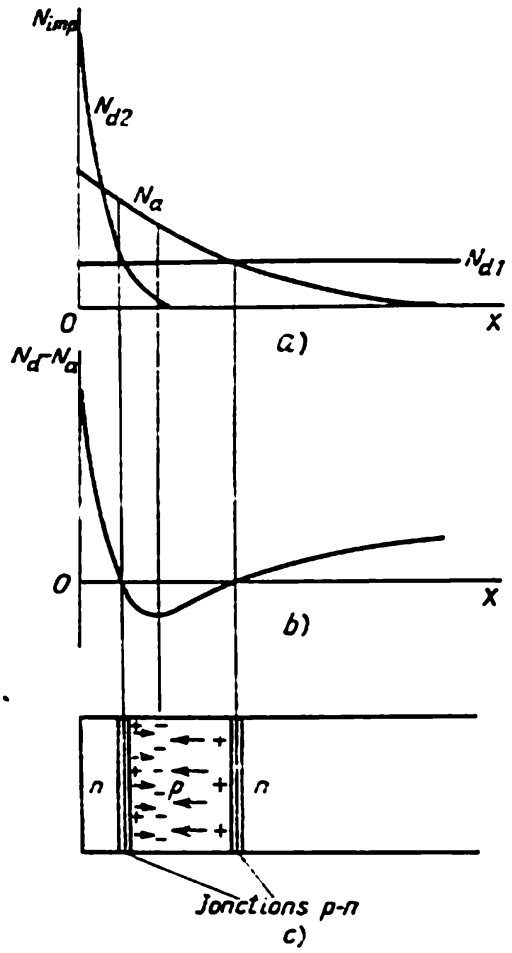

Fig. 12.8. Structure d'un transistor drift:

a — répartition de l'impureté dopante (x — distance à la surface; N_{d_1} — concentration initiale des donneurs; N_{d_2}, N_a — concentrations respectives des donneurs et des accepteurs ayant diffusé depuis la surface); *b* — répartition de l'impureté non compensée; *c* — répartition de la charge et champ « incorporé » (le sens du champ est indiqué par des flèches).

La durée des processus transitoires se détermine par le plus grand des temps considérés. Souvent, c'est le temps de diffusion des porteurs à travers la base (12.19). Pour réduire ce temps, il faut que la base soit aussi mince que possible. En plus de l'épaisseur w de la base la relation (12.19) fait intervenir encore le coefficient de diffusion des porteurs de courant minoritaires dans la base — des électrons dans un transistor *n-p-n* et des trous dans un transistor *p-n-p*. Le coefficient de diffusion des électrons étant plus grand que celui des trous (de deux fois pour le germanium et le silicium et de 20 fois pour le GaAs), les transistors *n-p-n* ont, toutes choses étant égales par ailleurs, des temps de commutation plus petits que les transistors *p-n-p*.

Les temps de commutation sont aussi réduits de façon substantielle dans des *transistors* dits *à dérive* ou *drift* dont la base possède un « champ incorporé » qui accélère le déplacement des porteurs minoritaires de l'émetteur vers le collecteur. De même que dans des diodes impulsionnelles (v. § 11.4), un tel champ est engendré par une impureté dopante inégalement répartie dans la base du transistor. Les régions d'émetteur et de base d'un transistor à dérive sont créées par la diffusion des impuretés donatrice et acceptrice dans une plaquette de type *n* (transistor *n-p-n*), comme l'indique la figure 12.8, *a*. La figure 12.8, *b* représente la répartition de l'impureté non compen-

sée dans la structure d'un transistor à dérive. La diffusion des trous dans la base depuis des régions à concentration augmentée en accepteurs non compensés produit des charges d'espace représentées sur la figure 12.8, *c*. Il est facile de voir que de telles charges engendrent un champ « incorporé » dans la base qui freine les électrons injectés de l'émetteur dans une petite région adjacente à l'émetteur, mais accélère leur mouvement vers le collecteur dans la partie principale de la région de base.

Comme caractéristique quantitative des propriétés fréquentielles d'un transistor, on prend une *fréquence limite d'amplification en*

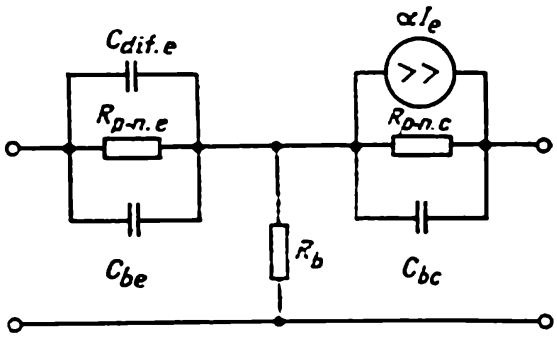

Fig. 12.9. Schéma équivalent d'un transistor.

courant ν_α, qu'on appelle fréquence de coupure alpha, pour laquelle le carré du module du gain en courant est deux fois plus faible que le carré du gain en courant $|\alpha_0|^2$ en basse fréquence: $|\alpha|^2/|\alpha_0|^2 = 0{,}5$. Les calculs montrent que

$$\nu_\alpha = \sqrt{3}/(2\pi t_{dif}), \tag{12.22}$$

où t_{dif} est donné par l'expression (12.19).

La technologie planar de diffusion utilisée par la fabrication des transistors dans laquelle les processus de diffusion sont effectués d'un seul côté de la plaquette, permet de réaliser des structures avec de très petites épaisseurs, connues avec précision, de la région de base. C'est pourquoi, dans les transistors drifts pour hautes fréquences modernes le temps de transit des porteurs dans la base est si petit que ce sont les temps de recharge des capacités de barrière des jonctions *p-n* émettrice et collectrice qui déterminent la durée des processus transitoires.

Le schéma équivalent le plus simple d'un transistor est représenté sur la figure 12.9. Le processus d'accumulation des porteurs dans la base du transistor, qui détermine le temps t_{dif} de leur diffusion à travers la base, est simulé par la capacité de diffusion de la jonction émettrice $C_{dif.e}$. Le fait que les porteurs minoritaires dans la base ayant atteint la jonction collectrice sont sollicités par son champ et transférés dans le collecteur est reflété par le branchement

d'un générateur de courant αI_c en parallèle avec la jonction *p-n* collectrice. Les résistances des jonctions *p-n* émettrice et collectrice sont représentées par $R_{p-n.e}$ et $R_{p-n.c}$ respectivement et la résistance de la région de base par R_b, les capacités de barrière des jonctions *p-n* par les condensateurs $C_{b.e}$ et $C_{b.c}$.

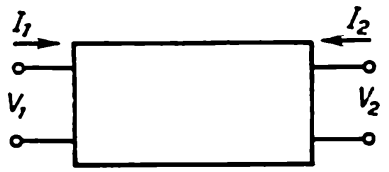

Fig. 12.10. Représentation d'un transistor par un quadripôle.

Les éléments constitutifs du circuit équivalent d'un transistor ne sont pas linéaires ; les résistances et les réactances varient avec les tensions auxquelles elles sont soumises et les courants qui les parcourent. Aussi, un tel schéma n'est-il commode que pour des signaux de faible valeur, lorsque les courants et les tensions sont liés par des relations linéaires. Dans de tels cas un transistor peut être considéré de façon formelle comme un quadripôle linéaire (fig. 12.10). Il existe plusieurs systèmes de paramètres d'un quadripôle qui lient entre elles les valeurs de faibles tensions et courants variables à l'entrée ($\tilde{V}_1$ et $\tilde{I}_1$) et la tension et le courant à la sortie ($\tilde{V}_2$ et $\tilde{I}_2$). Les plus employés sont les *paramètres hybrides h* :

$$\begin{cases} \tilde{V}_1 = h_{11}\tilde{I}_1 + h_{12}\tilde{V}_2, \\ \tilde{I}_2 = h_{21}\tilde{I}_1 + h_{22}\tilde{V}_2. \end{cases} \tag{12.23}$$

Pour déterminer le paramètre h_{11} il faut mesurer la tension variable et le courant variable à l'entrée de transistor, pour une tension variable $\tilde{V}_2 = 0$ à sa sortie. Cette condition est facile à accomplir en court-circuitant la sortie par un condensateur de capacité suffisamment grande. Pour déterminer la valeur de h_{12} il convient d'appliquer un signal variable $\tilde{V}_2$ aux bornes de sortie du quadripôle et de mesurer la tension $\tilde{V}_1$ entre ses bornes d'entrée, en utilisant à cet effet un voltmètre présentant une résistance interne suffisamment grande pour qu'on puisse considérer que $\tilde{I}_1 = 0$. Le paramètre h_{12} constitue ainsi le coefficient de réaction en tension. Enfin, $h_{21} = \tilde{I}_2/\tilde{I}_1$ pour $\tilde{V}_2 = 0$ et $h_{22} = \tilde{I}_2/\tilde{V}_2$ pour $\tilde{I}_1 = 0$.

Outre les paramètres *h* on emploie aussi les *paramètres z* et *y* :

$$\begin{cases} \tilde{V}_1 = z_{11}\tilde{I}_1 + z_{12}\tilde{I}_2, \\ \tilde{V}_2 = z_{21}\tilde{I}_1 + z_{22}\tilde{I}_2 ; \end{cases} \tag{12.24}$$

$$\begin{cases} \tilde{I}_1 = y_{11}\tilde{V}_1 + y_{12}\tilde{V}_2, \\ \tilde{I}_2 = y_{21}\tilde{V}_1 + y_{22}\tilde{V}_2. \end{cases} \tag{12.25}$$

Les relations entre les divers systèmes de paramètres sont faciles à établir. C'est ainsi qu'en exprimant $\tilde{V}_1$ figurant dans (12.23) moyennant $\tilde{I}_1$ et $\tilde{I}_2$:

$$\tilde{V}_1 = \left(h_{11} - \frac{h_{12}h_{21}}{h_{22}}\right)\tilde{I}_1 + \frac{h_{12}}{h_{22}}\tilde{I}_2, \tag{12.26}$$

et en comparant l'expression ainsi obtenue avec la première équation de (12.24), on trouve

$$z_{11} = h_{11} - \frac{h_{12}h_{21}}{h_{22}}; \quad z_{12} = \frac{h_{12}}{h_{22}}. \tag{12.27}$$

Il convient de signaler que les paramètres d'un quadripôle dépendent du montage du transistor — à base commune ou à émetteur commun. Le lien entre les paramètres de ces deux montages peut être facilement établi si l'on tient compte que

$$\begin{cases} \tilde{I}_e + \tilde{I}_b + \tilde{I}_c = 0, \\ \tilde{V}_{c.e} = \tilde{V}_{c.b} - \tilde{V}_{e.b}. \end{cases} \tag{12.28}$$

Dans les expressions (12.28) les courants sont comptés positivement lorsqu'ils entrent dans la région de base. Pour le montage à base commune, $\tilde{I}_1 = \tilde{I}_e$, $\tilde{V}_1 = \tilde{V}_{e.b}$, $\tilde{I}_2 = \tilde{I}_c$, $\tilde{V}_2 = \tilde{V}_{c.b}$; pour le montage à émetteur commun $\tilde{I}_1 = \tilde{I}_b$, $\tilde{V}_1 = -\tilde{V}_{e.b}$, $\tilde{I}_2 = \tilde{I}_c$, $\tilde{V}_2 = \tilde{V}_{c.e}$. Ainsi, pour passer par exemple des paramètres h du montage à base commune aux paramètres h du montage à émetteur commun il faut exprimer $\tilde{I}_e$ en fonction de $\tilde{I}_b$ et $\tilde{I}_c$ et $\tilde{V}_{c.b}$ en fonction de $\tilde{V}_{c.e}$, en utilisant à cet effet les relations (12.28), introduire les grandeurs obtenues dans (12.23) et mettre cette dernière sous la forme

$$\begin{cases} -\tilde{V}_{e.b} = h_{11}^e \tilde{I}_b + h_{12}^e \tilde{V}_{c.e}, \\ \tilde{I}_c = h_{21}^e \tilde{I}_b + h_{22}^e \tilde{V}_{c.e}, \end{cases} \tag{12.29}$$

où h_{ik}^e désignent les paramètres k du montage à émetteur commun.

§ 12.4. Phototransistors

Un transistor peut être utilisé comme photorécepteur à *amplification interne* du courant photoélectrique primaire. Un *phototransistor* est un transistor ordinaire dont la région de base peut être éclairée (fig. 12.11, *a*). La figure 12.11, *b* montre le montage le plus répandu d'un transistor à émetteur commun et à base libre (souvent les phototransistors n'ont pas du tout de sortie de base). En l'absence de flux lumineux W_0, le transistor est parcouru par un courant d'obscurité.

Lorsque la région de base d'un phototransistor du type *n-p-n* par exemple est soumise à l'éclairement avec une énergie de quanta $h\nu > E_g$, dans la base il y a création de paires électron-trou. En diffusant dans la région de base, ces paires atteignent la jonction collectrice où elles sont annulées par le champ de cette jonction. Les électrons sont poussés vers la région de collecteur en produisant un *photocourant* dit *primaire* tandis que les trous restent dans la base où ils forment une charge d'espace positive. La présence de cette charge

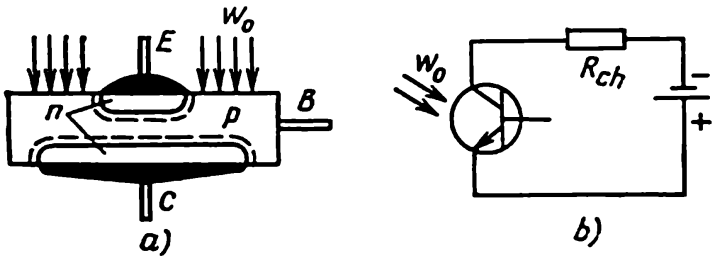

Fig. 12.11. Constitution (*a*) et schéma de branchement (*b*) d'un phototransistor.

a pour effet de réduire la hauteur de la barrière de potentiel de la jonction *p-n* émettrice de sorte que les électrons sont injectés dans la base où ils compensent de nouveau la charge due aux trous. Etant animés d'une agitation thermique désordonnée, c'est-à-dire en diffusant dans la base, les électrons « tombent » dans la région de collecteur, de nouveaux électrons en provenance de l'émetteur viennent occuper leurs places et ainsi de suite tant que les trous excédentaires produits dans la base par la lumière ne se recombinent avec les électrons. Ainsi, pendant la durée de vie plusieurs électrons, par chaque trou, passeront par la jonction collectrice — il y a amplification du photocourant primaire.

Le courant de collecteur I_c est la somme du courant de génération thermique I_{sc}, du photocourant primaire I_{ph} et du courant d'électrons αI_e injectés à travers la jonction émettrice. En même temps, dans le montage à base ouverte $I_c = I_e = I$. Ainsi,

$$I = I_{sc} + I_{ph} + \alpha I. \tag{12.30}$$

D'où

$$I = (I_{sc} + I_{ph})/(1 - \alpha). \tag{12.31}$$

Les courants $I_{sc}/(1 - \alpha)$ et $I_{ph}/(1 - \alpha)$ sont respectivement le courant d'obscurité et le photocourant primaire amplifié $1/(1 - \alpha)$ fois. Si $\alpha = 0{,}995$, $1/(1 - \alpha) = 200$. Ainsi, la sensibilité d'un phototransistor peut être très supérieure à celle d'une photodiode dont le signal est égal tout simplement à I_{ph}. Il est toutefois à noter que ce gain en sensibilité s'accompagne d'une perte en rapidité de fonctionnement. En effet, le temps de photoréponse d'un phototransistor se détermine par la durée de vie τ des porteurs hors d'équilibre dans sa

région de base alors que dans une photodiode il se détermine par le temps de diffusion des porteurs de la région, où ils sont générés par la lumière, jusqu'à la jonction *p-n* qui est beaucoup plus faible que τ ou seulement par le temps de leur transit dans la région de charge d'espace de la jonction *p-n* si la photogénération se produit principalement dans la jonction *p-n* elle-même.

§ 12.5. Transistors à effet de champ

Ces dernières années une grande place dans l'électronique a été occupée par des appareils utilisant des phénomènes qui interviennent dans la couche de surface d'un semiconducteur. L'élément fondamental de tels appareils est constitué par une structure métal-diélectrique-semiconducteur (MDS) (fig. 12.12). Souvent le diélectrique est constitué par une couche d'oxyde, par exemple de SiO_2. De telles

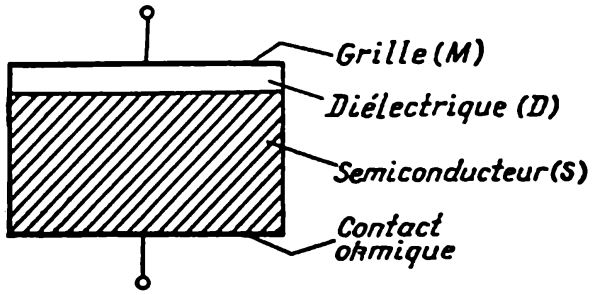

Fig. 12.12. Schéma d'une structure MOS.

structures portent le nom de *structures MOS*. L'électrode métallique est déposée sur le diélectrique le plus souvent par évaporation sous vide. Cette électrode est appelée *grille* (*gate* dans la littérature anglaise et américaine).

Si l'on applique à la grille une certaine tension de polarisation par rapport au semiconducteur, il apparaît, près de la surface du semiconducteur, une région de charge d'espace de signe contraire à celui de la charge de la grille. Dans cette région la concentration des porteurs de courant peut différer de façon substantielle de leur concentration à l'intérieur du semiconducteur.

La charge de la région de surface du semiconducteur a pour effet de provoquer l'apparition d'une différence de potentiel entre cette région et les régions profondes du semiconducteur et donc une courbure des bandes d'énergie. Si la charge de la grille est négative, les bandes d'énergie s'incurvent vers le haut parce que le déplacement d'un électron des régions profondes vers la surface augmente son énergie potentielle (fig. 12.13). Si la grille est chargée positivement, les bandes d'énergie s'incurvent vers le bas.

La figure 12.13 montre la structure de bande d'un semiconducteur de type *n* pour le cas où la grille est chargée négativement et indi-

que les désignations des grandeurs principales caractérisant la surface : ψ_s — différence de potentiel entre la surface et les régions profondes du semiconducteur ; $\varphi_s = -e\psi_s$ — courbure des bandes d'énergie près de la surface ; E_i — milieu de la bande interdite. Comme il est visible sur la figure 12.13, la distance entre le niveau inférieur de la bande de conduction et le niveau de Fermi ($-\mu$) est

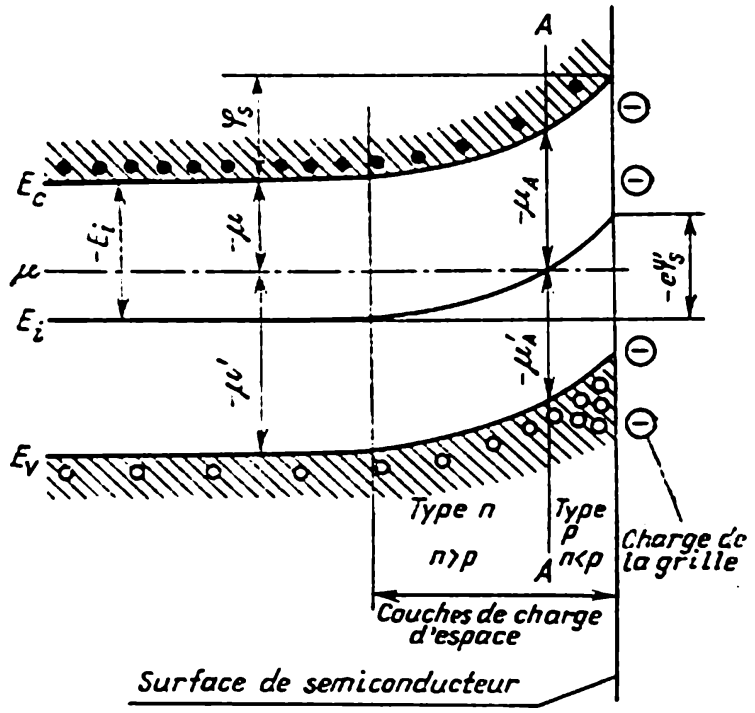

Fig. 12.13. Courbure des bandes près de la surface d'un semiconducteur de type n lorsque la grille est chargée négativement.

en volume du semiconducteur plus petite que la distance entre le niveau de Fermi et le sommet de la bande de valence ($-\mu'$). C'est pourquoi la concentration d'équilibre des électrons n_0 est plus forte que celle des trous : $n_0 > p_0$, comme il se doit pour un semiconducteur de type n. Dans la couche de surface de la charge d'espace il se produit une courbure des bandes d'énergie et la distance entre le niveau inférieur de la bande de conduction et le niveau de Fermi augmente de façon continue à mesure que l'on s'approche de la surface alors que la distance entre le niveau de Fermi et le sommet de la bande de valence diminue de façon continue. Dans une section telle que AA ces distances s'égalisent ($-\mu_A = -\mu'_A$) et le semiconducteur devient intrinsèque : $n=p=n_i$. A droite de la section AA on a $|\mu| > |\mu'|$, de sorte que $p > n$ et le semiconducteur devient de type p. Près de la surface il se forme dans ce cas une jonction p-n superficielle.

La courbure des bandes d'énergie près de la surface est souvent ex-

primée en unités de kT et notée Y_s. Alors

$$Y_s = e\psi_s/(kT) = -\varphi_s/(kT). \tag{12.32}$$

Lors de la formation de la région de surface d'un semiconducteur trois cas importants peuvent se présenter: l'*appauvrissement*, l'*in-*

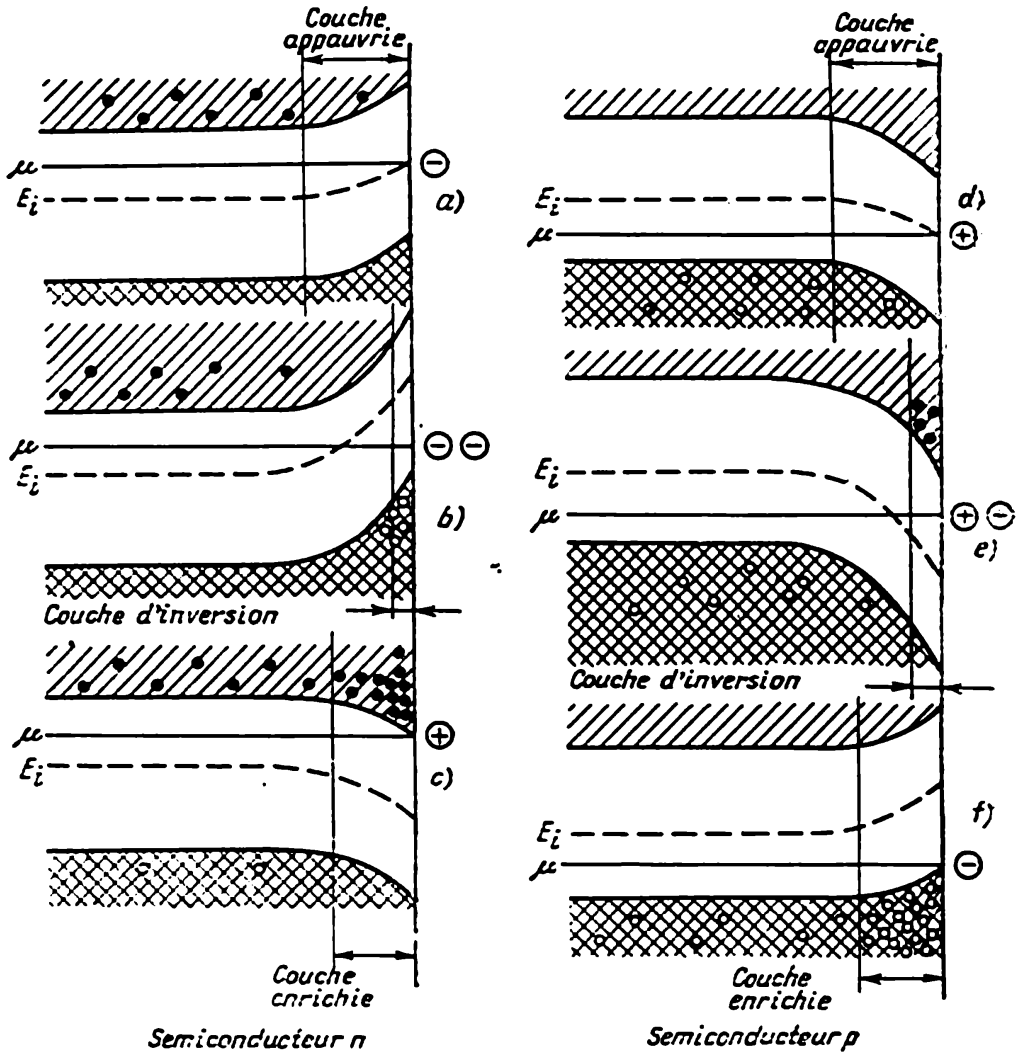

Fig. 12.14. Formation d'une couche superficielle de charge d'espace au-dessous de la grille dans les semiconducteurs des types n et p.

version et l'*enrichissement* de cette région par les porteurs de charges. Ces cas sont représentés pour des semiconducteurs des types n et p sur la figure 12.14.

Une région appauvrie apparaît dans le cas où le signe de la charge de grille est le même que celui des porteurs de courant majoritaires

(fig. 12.14, *a*, *d*). La courbure des bandes d'énergie provoquée par une telle charge conduit à l'augmentation de la distance séparant le niveau de Fermi du niveau inférieur de la bande de conduction dans un semiconducteur de type *n* et le niveau de Fermi du sommet de la bande de valence dans un semiconducteur de type *p*. L'augmentation de cette distance s'accompagne de l'appauvrissement de la région de surface en porteurs majoritaires : la charge de la grille chasse les porteurs majoritaires de la région de surface.

Pour une forte densité de charge de la grille dont le signe est le même que celui de la charge des porteurs majoritaires, à mesure que l'on s'approche de la surface, la distance entre le niveau de Fermi et le sommet de la bande de valence dans un semiconducteur de type *n* devient inférieure à la distance séparant le niveau de Fermi du niveau inférieur de la bande de conduction. Il en résulte que la concentration des porteurs minoritaires (des trous) près de la surface du semiconducteur devient supérieure à la concentration des porteurs majoritaires et le type de conductibilité de cette région change bien que les électrons et les trous y soient peu nombreux presque comme dans un semiconducteur intrinsèque. Mais près de la surface elle-même, les porteurs minoritaires peuvent être aussi nombreux et même plus nombreux que les porteurs majoritaires dans la masse du semiconducteur. De telles couches superficielles, bons conducteurs, dont le type de conductibilité est inverse de celui en profondeur, sont appelées *couches d'inversion* (elles sont montrées sur la figure 12.14, *b*, *e*). A la couche d'inversion est adjacente une couche de déplétion.

Si le signe de la charge de la grille est contraire au signe de la charge des porteurs majoritaires dans le semiconducteur, les porteurs majoritaires sont attirés sous l'effet de cette charge vers la surface, ce qui enrichit la couche superficielle en porteurs majoritaires. De telles couches sont dites *enrichies* (elles sont montrées sur la figure 12.14, *c*, *f*).

Dans le cas d'un faible enrichissement ou appauvrissement de la région superficielle d'un semiconducteur sa dimension se détermine par la longueur dite de Debye :

$$L_D = \sqrt{\frac{\varepsilon\varepsilon_0 kT}{e^2 n_0}}, \qquad (12.33)$$

où ε est la permittivité du semiconducteur ; n_0 la concentration des porteurs de courant majoritaires dans le semiconducteur. Sur la longueur L_D de la couche l'intensité de champ électrique, le potentiel et la variation de la concentration des porteurs de courant par rapport aux régions profondes du semiconducteur tombent de *e* fois par rapport à leurs valeurs de surface.

Dans le cas d'un appauvrissement fort, l'épaisseur de la couche appauvrie peut être calculée à l'aide de la formule pour la couche de déplétion dans la barrière de Schottky (10.18) en y remplaçant φ_0

par φ_s. C'est le calcul de la structure de la couche superficielle à région d'inversion qui est un des plus complexes.

Dans l'électronique des circuits intégrés on emploie couramment les structures MOS pour la réalisation des transistors et, à leur base, de divers microcircuits intégrés. La figure 12.15 représente schématiquement la structure d'un transistor MOS à grille isolée. Ce transistor est constitué par un cristal de silicium (de type n par exemple)

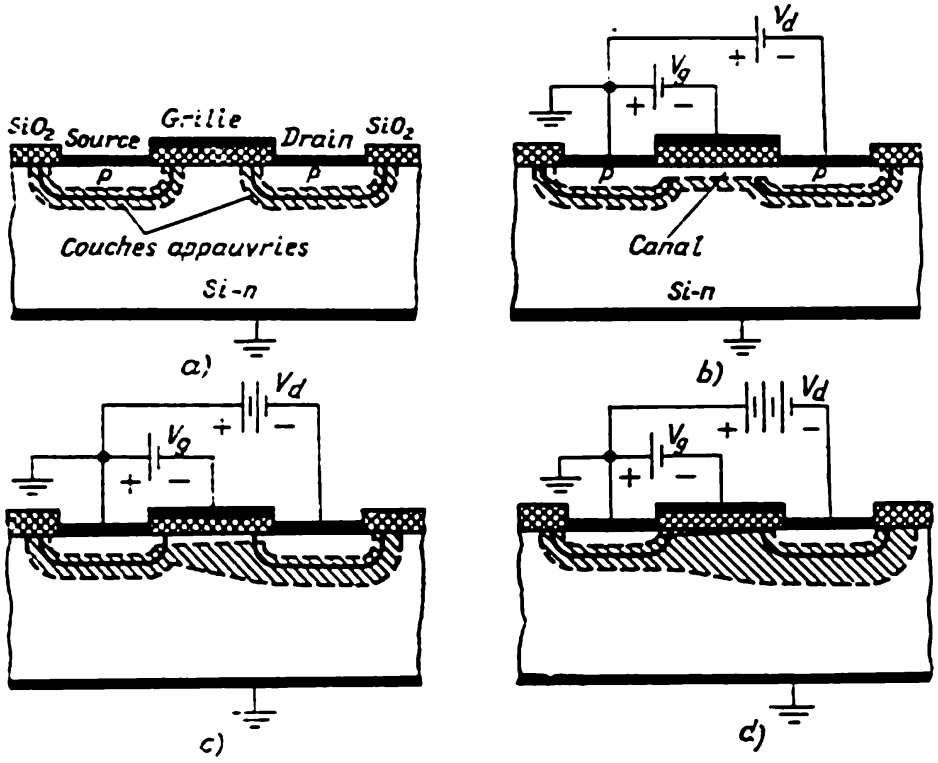

Fig. 12.15. Explication du fonctionnement d'un transistor MOS:
a — structure d'un transistor MOS; b — formation d'un canal entre la source et le drain; c — pincement du canal en cas d'augmentation de la tension entre la source et le drain; d — recouvrement du canal.

près de la surface duquel sont formées, par diffusion (ou par implantation ionique) à travers des fenêtres pratiquées dans la couche d'oxyde, des régions p comme l'indique la figure 12.15, *a*. Une de ces régions est appelée *source* et l'autre *drain*. En haut, ces régions portent des contacts ohmiques. L'intervalle entre les régions est recouvert d'un film métallique isolé de la surface du cristal par une couche d'oxyde. Comme il a déjà été dit, cette électrode du transistor porte le nom de grille. A la surface frontière entre les régions p et n il se forme deux jonctions p-n: de source et de drain qui sont couvertes de hachures sur la figure 12.15, *a*.

La figure 12.15, *b* montre le schéma de branchement du transistor dans le circuit: le pôle positif de la source de tension V_d est relié à la source et le pôle négatif au drain, le pôle négatif de la sour-

ce de tension V_g est relié à la grille. Pour simplifier l'exposé, supposons que la différence de potentiel de contact, la charge dans l'oxyde et les états de surface (v. plus loin) n'existent pas. Alors, en l'absence de la tension sur la grille, les propriétés de la région superficielle ne diffèrent en rien des propriétés des régions profondes du semiconducteur. La résistance entre le drain et la source est très grande parce que la jonction *p-n* de drain est polarisée en sens inverse. L'application à la grille d'une tension de polarisation négative conduit d'abord à la formation, au-dessous de la grille, d'une région appauvrie et ensuite, pour une certaine tension dite *de seuil* $V_{g.seuil}$, à la formation d'une région d'inversion qui relie entre elles les régions *p* de source et de drain par un *canal conducteur*. Lorsque les tensions agissant sur la grille deviennent supérieures à la valeur de seuil $V_{g.seuil}$, le canal s'élargit et la résistance drain-source diminue. La structure que nous considérons est donc une *résistance contrôlable*.

Pourtant la résistance du canal ne se détermine par la tension de grille que pour de petites tensions de drain. Lorsque V_d augmente, les porteurs s'en vont du canal dans la région de drain, la couche appauvrie près de la jonction *p-n* de drain s'élargit et le canal se rétrécit (fig. 12.15, *c*). La variation du courant I_d en fonction de la tension de drain V_d devient non linéaire.

Lorsque le canal se rétrécit, le nombre de porteurs de courant libres au-dessous de la grille diminue à mesure que l'on s'approche du drain. Pour que le courant dans le canal soit le même dans toute section, le champ électrique le long du canal doit être dans ce cas non uniforme, son intensité doit augmenter à mesure que l'on s'approche du drain. En outre, l'apparition du gradient de concentration des porteurs de courant libres le long du canal a pour effet de provoquer l'apparition d'une composante de diffusion de la densité de courant.

Pour une certaine tension de drain $V_{d.p}$ le canal est pincé près du drain et pour une polarisation encore plus grande le canal se rétrécit vers la source (fig. 12.15, *d*). Toutefois, l'étranglement du canal ne fait pas disparaître le courant de drain parce que dans la couche appauvrie qui coupe le canal, le champ électrique tire les trous le long de la surface. Lorsque les porteurs de courant pénètrent par diffusion du canal dans cette région, ils sont entraînés par le champ qui les pousse vers le drain. Ainsi à mesure que la tension agissant sur le drain augmente, le mécanisme de déplacement uniquement par dérive des porteurs de courant le long du canal est remplacé par le mécanisme de déplacement par dérive et diffusion.

Le mécanisme de passage du courant dans un transistor MOS dans le cas du canal pincé présente certains traits communs avec le passage du courant à travers une jonction *p-n* polarisée en sens inverse. Rappelons que dans une jonction *p-n* les porteurs de courant minoritaires pénètrent par diffusion dans la région de charge d'espace où ils sont sollicités par le champ de cette jonction.

Comme le montrent la théorie et l'expérience, après le pincement du canal le courant de drain I_d atteint pratiquement sa valeur de saturation (fig. 12.16). La valeur du courant de saturation dépend de la tension de grille V_g : plus élevée est cette tension, plus large est le canal et plus fort est le courant de saturation (sur la figure $|V_{g3}| > |V_{g2}| > |V_{g1}|$). C'est un effet typiquement transistor — en agissant sur la tension de grille (tension dans le circuit d'entrée), on peut commander le courant de drain (courant dans le circuit de sortie). Une particularité caractéristique d'un transistor MOS est que son entrée peut être constituée par un condensateur représenté par la grille métallique isolée du semiconducteur. Les courants de fuite de grille dans les transistors MOS typiques sont de l'ordre de 10^{-15} A et peuvent être réduits jusqu'à 10^{-17} A. Le transistor que nous considérons utilise l'effet de champ, ce qui explique l'appellation de transistors *à effet de champ*. A la différence des transistors du type *p-n-p* ou *n-p-n* dans lesquels les porteurs de courant minoritaires sont injectés dans la région de base, dans les transistors à effet de champ le courant n'est transporté que par les porteurs majoritaires. C'est la raison pour laquelle de tels transistors sont dits encore *unipolaires*.

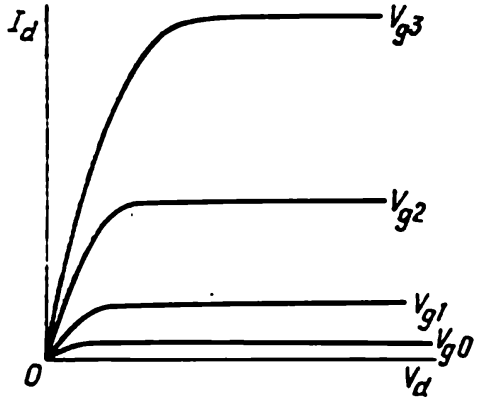

Fig. 12.16. Caractéristiques courant-tension d'un transistor MOS.

A la surface de séparation semiconducteur-diélectrique, dans la bande interdite du semiconducteur il existe des états énergétiques appelés *états de surface* ou plus exactement *états de surface de séparation* (fig. 12.17). Les fonctions d'onde des électrons dans ces états sont localisées près de la surface de séparation dans des régions de l'ordre de la constante du réseau cristallin. La cause de l'apparition de ces états est l'imperfection de la surface de séparation semiconducteur-diélectrique (oxyde). Les surfaces de séparation réelles comportent toujours un certain nombre de liaisons coupées et la stœchiométrie de la composition de la couche d'oxyde diélectrique est troublée. La densité et le caractère des états de la surface de séparation dépendent, pour une large part, de la technologie utilisée pour la réalisation de la couche diélectrique.

L'existence des états de surface à la limite de séparation semiconducteur-diélectrique a une influence négative sur les paramètres d'un transistor MOS parce qu'une partie de la charge induite au-dessous de la grille dans le semiconducteur est capturée par ces états. Un succès dans la réalisation des transistors du type considéré a

été obtenu après la mise au point de la technologie permettant de déposer sur la surface de silicium des films de SiO_2 avec une faible densité des états de la surface de séparation Si-SiO_2.

Dans le dioxyde de silicium lui-même il existe toujours une charge « incorporée » positive dont l'origine n'est pas encore claire. La valeur de cette charge dépend de la technologie de fabrication de l'oxyde et est souvent si grande que si le substrat utilisé est le silicium *p*, une couche d'inversion se forme près de sa surface même pour une polarisation nulle de la grille. De tels transistors sont dits *transistors MOS à canal dopé*. Leur canal subsiste même si la grille

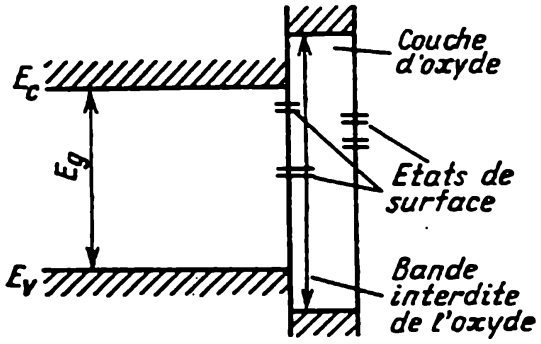

Fig. 12.17. Etats de surface.

est soumise à une certaine tension de polarisation négative. A la différence de ces derniers, dans les transistors réalisés sur un substrat *n* ou sur un substrat *p* fortement dopé où la formation d'une couche d'inversion exige une charge trop grande de l'oxyde, le canal ne se forme que si la grille est attaquée par une tension V_g supérieure à une certaine valeur de seuil $V_{g.seuil}$. Le signe de cette polarisation de la grille doit être négatif pour les transistors à substrat *n* et positif pour les transistors à substrat *p*.

Dans la catégorie des transistors unipolaires on range aussi des transistors à jonction *p-n* de commande dont la structure est représentée schématiquement sur la figure 12.18, *a*. Dans de tels transistors le canal conducteur représente une région étroite dans le semiconducteur de départ non occupée par la couche appauvrie de la jonction *p-n*. On peut contrôler la largeur de cette région en polarisant la jonction *p-n* en sens inverse. Cette polarisation fait varier la résistance initiale drain-source. Si, la tension sur la jonction *p-n* étant constante, on augmente la polarisation appliquée entre le drain et la source, le canal se rétrécit vers le drain et le courant de drain croît avec la tension plus lentement que dans le cas des polarisations faibles. Au pincement du canal (fig. 12.18, *b*) le courant de drain atteint sa valeur de saturation. Aussi bien le mécanisme du passage du courant par le canal d'un tel transistor, que ses caractéristiques de sortie sont assez proches des caractéristiques correspondantes d'un transistor MOS.

La réactance d'entrée des transistors à effet de champ aux basses fréquences est purement capacitive. La capacité d'entrée $C_{g\text{-}s}$ est

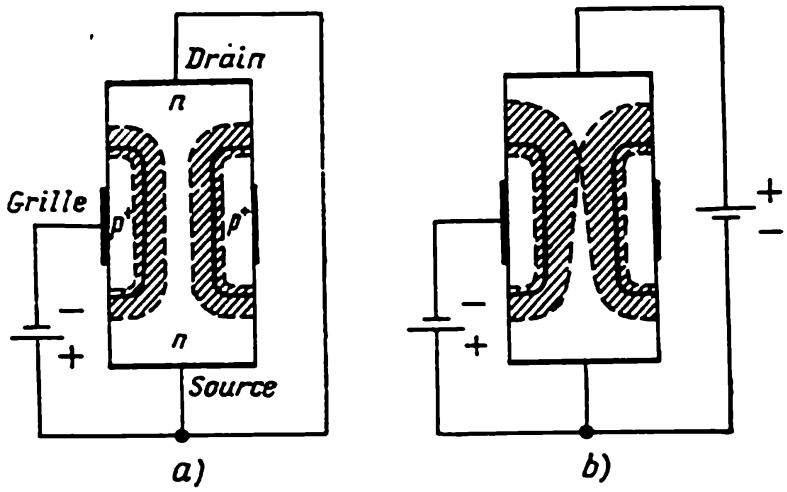

Fig. 12.18. Explication du principe de fonctionnement d'un transistor MOS à canal, à jonction *p-n* de commande:
a — polarisation nulle du drain; *b* — recouvrement du canal en cas d'augmentation de la polarisation du drain.

constituée par la grille et par la partie non recouverte du canal du côté de la source. Vu que pour la charge de cette capacité le courant doit passer par la partie non pincée du canal de résistance R_c, la constante de temps propre du transistor est égale à $C_{g\text{-}s}R_c$. Toutefois, ce temps est très petit de sorte que dans les circuits intégrés équipant par exemple des ordinateurs digitaux la durée des processus transitoires est déterminée non par ce temps mais par des capacités parasites du montage et par des capacités d'entrée des autres transistors branchés sur la sortie du transistor considéré. Ceci étant, lors de la réalisation de tels circuits on cherche à diminuer autant que possible la capacité d'entrée grâce à la réduction de la longueur du canal et à une coïncidence rigoureuse des frontières de la grille avec les frontières de la source et du drain.

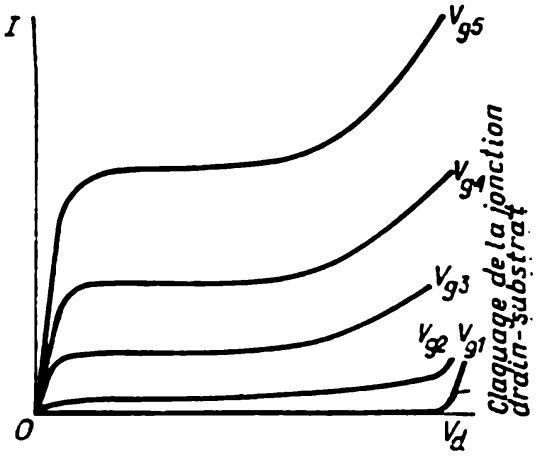

Fig. 12.19. Caractéristiques de claquage d'un transistor MOS classique.

Dans le cas des fortes tensions sur le drain d'un transistor MOS, la région de charge d'espace peut se propager depuis la région de drain si fortement que le canal disparaît totalement. Dans un tel cas,

les porteurs de la région de source fortement dopée se précipiteront sur le drain exactement de même qu'en cas de « percement » de la base d'un transistor bipolaire (v. § 12.2 et fig. 12.6).

Si la tension sur le drain est suffisamment grande, il peut se produire aussi un claquage ordinaire de la jonction *p-n* de drain polarisée en sens inverse. Les caractéristiques de sortie d'un transistor MOS présentant des parties de claquage sont montrées sur la figure 12.19.

§ 12.6. Appareils à caractéristique courant-tension en S

Diodes à base longue. Dans les dispositifs d'automatisme et la technique de calcul électronique une place importante est accordée aux appareils à caractéristique courant-tension en S (fig. 12.20). Si les tensions de polarisation V_{pol} et la résistance de charge (elle détermine la pente des droites de la figure 12.20) sont réglées à des valeurs convenables, le circuit utilisant un tel appareil possède deux états stables (*1* et *3* sur la figure 12.20; on peut montrer que l'état *2* n'est pas stable). Dans l'état *1*, la partie principale de la tension de polarisation chute dans la diode et le circuit est parcouru par un courant faible I_1, dans l'état *3*, la partie principale de la polarisation chute dans la charge et le circuit est parcouru par un courant fort I_3. Le passage de la diode S de l'état à haute résistance (*1*) à l'état à résistance faible (*3*) se produit lorsque la tension de polarisation devient égale à la tension de commutation V_{com} comme l'indique la flèche de la figure 12.20.

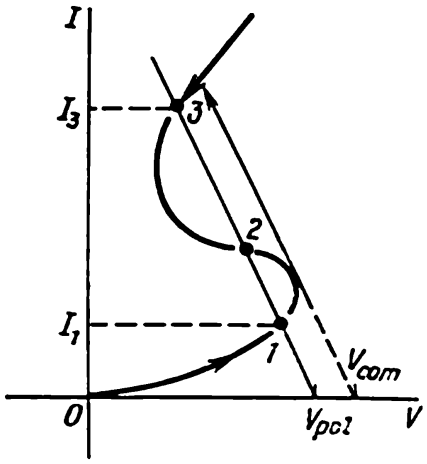

Fig. 12.20. Caractéristique courant-tension en S.

Ainsi, l'appareil à caractéristique courant-tension en S est un commutateur électronique.

Les appareils les plus simples qui peuvent avoir une caractéristique courant-tension en S sont des *diodes à base longue* (l'épaisseur de la région à forte résistance dépasse dans de telles diodes la longueur de diffusion des porteurs minoritaires dans cette région). Les porteurs minoritaires injectés à travers une jonction *p-n* polarisée en sens direct diminuent la résistance des régions *p* et *n* de la diode, ce qui favorise une redistribution de la tension appliquée à la diode au profit de la jonction *p-n*. L'augmentation de la tension agissant sur la jonction *p-n* conduit à l'augmentation du courant d'injection et donc

à une diminution encore plus grande de la résistance série et ainsi de suite. L'analyse montre pourtant qu'une partie à résistance différentielle négative apparaît sur la caractéristique courant-tension de la diode dans le cas où la conductance de la région de base de la diode croît non pas proportionnellement au courant d'injection, mais plus brusquement.

Il existe plusieurs mécanismes qui peuvent conduire à une croissance plus forte que linéaire de la conductance en fonction du courant d'injection. Un d'eux est l'augmentation de la durée de vie

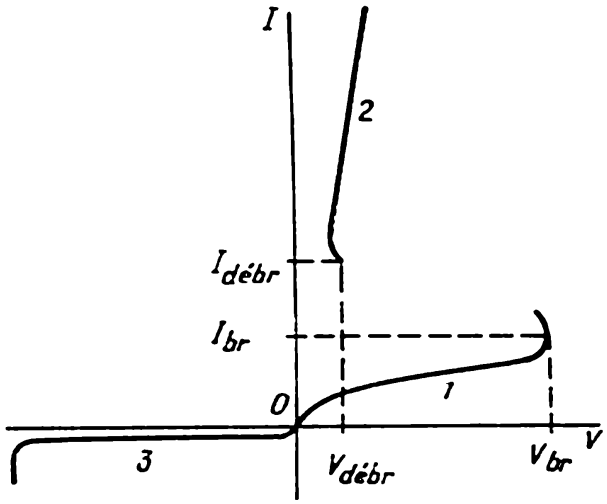

Fig. 12.21. Caractéristique courant-tension d'un dynistor.

des porteurs hors d'équilibre avec l'augmentation de leur concentration lors de la recombinaison par le niveau des pièges (v. § 8.2). Une augmentation supplémentaire de la conductibilité de la région de base s'obtient aussi lorsqu'elle absorbe le rayonnement bande à bande émis lors de la recombinaison des porteurs hors d'équilibre dans les matériaux semiconducteurs à bandes droites du type de l'arséniure de gallium.

Les plus employés parmi les éléments à caractéristique courant-tension en S sont toutefois des structures semiconductrices multicouches — les *diodes à trijonction* (*diodes à quatre régions*) et les *thyristors*.

Diodes à trijonction. La caractéristique courant-tension d'une diode à jonction *p-n-p-n* (appelée encore thyristor non commandé ou dynistor) est représentée par la fig. 12.21 et sa structure par la fig. 12.22, *a*. L'appareil comporte trois jonctions *p-n*. Les jonctions latérales sont dites *émettrices* et la jonction centrale est *collectrice*. Les régions médianes de la structure, appelées *bases*, sont dopées assez faiblement. Les régions latérales (émettrices) ont une concentration d'impuretés plus forte. Pour la polarisation en sens direct le pôle positif de la source de tension est relié à la région p_1 latérale et le pôle négatif à la région n_2. Dans ce cas, les jonctions p_1-n_1 et p_2-n_2 se trouvent

montées en sens direct et la jonction centrale n_1-p_2 en sens inverse (fig. 12.22, *b*). La résistance de la jonction n_1-p_2 polarisée en sens inverse étant grande, toute la tension appliquée à l'appareil chute pratiquement dans cette jonction et la polarisation directe des jonctions p_1-n_1 et p_2-n_2 est peu élevée. Par conséquent, le courant traversant l'appareil est lui aussi peu intense. Déterminons sa valeur. Cela peut se faire le plus commodément en considérant le courant I_c à travers la jonction n_1-p_2.

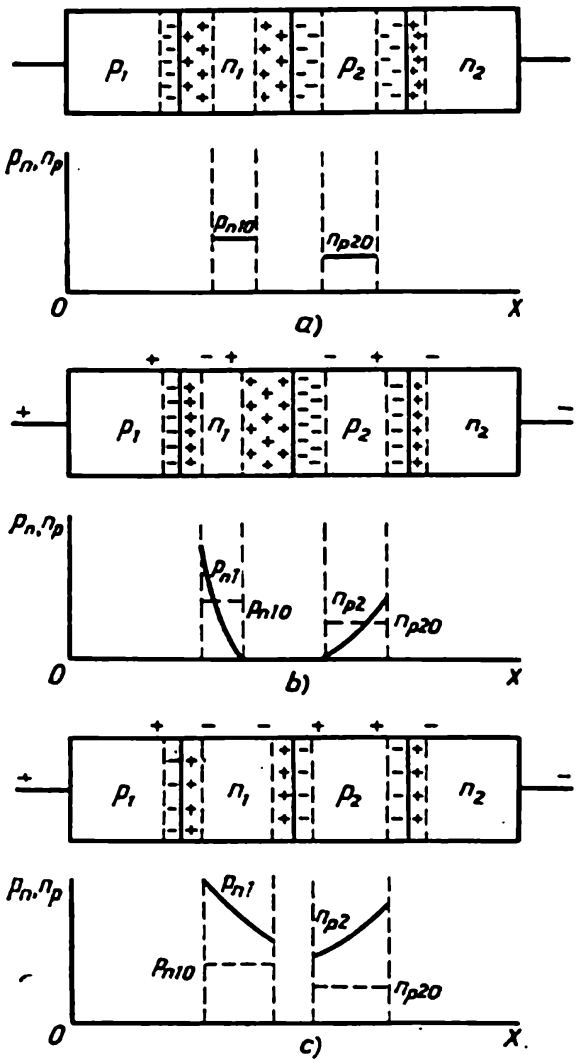

Fig. 12.22. Répartition des charges d'espace et des concentrations des porteurs minoritaires dans les bases d'un thyristor à l'état d'équilibre (*a*), au repos (*b*) et en marche de régime (*c*).

Comme cette jonction est polarisée en sens inverse, les porteurs de courant majoritaires ne peuvent pas franchir sa haute barrière de potentiel et le courant est transporté à travers la jonction par des porteurs minoritaires. Les porteurs de courant minoritaires sont produits par suite de la génération thermique d'une part et sont injectés dans les régions n_1 et p_2 à travers les jonctions p_1-n_1 et p_2-n_2 (ces jonctions sont polarisées en sens direct) d'autre part. Une partie considérable des porteurs minoritaires injectés atteignent la jonction collectrice du fait que les régions n_1 et p_2 sont réalisées plus minces que les longueurs de diffusion des porteurs minoritaires L_p et L_n respectivement.

Ainsi, le courant de jonction collectrice I_c est la somme du courant de thermogénération I_{cs}, du courant de trous I_{cp} et du courant d'électrons I_{cn} injectés à travers les jonctions émettrices et ayant atteint le collecteur. Les deux derniers courants peuvent être déterminés en partant des considérations suivantes. Le courant de trous injectés à travers la jonction p_1-n_1 constitue une partie du courant total I_1 passant par cette jonction. Il est égal à $\gamma_1 I_1$, où γ_1 est le *coefficient d'injection* de la jonction émettrice. Le courant d'électrons traversant la jonction p_2-n_2 est égal à $\gamma_2 I_2$, où I_2 est le courant total

de la jonction p_2-n_2 et γ_2 le coefficient d'injection de cette jonction. Le courant de trous I_{cp} qui ont été injectés à travers la jonction p_1-n_1 et ont atteint le collecteur sans se recombiner est égal à $\beta_1\gamma_1 I_1$. Ici, β_1 est le *coefficient de transfert* des trous à travers la région n_1. D'une manière analogue, le courant d'électrons I_{cn} qui ont atteint, en passant par la région p_2, la jonction collectrice est égal à $\beta_2\gamma_2 I_2$, où β_2 est le coefficient de transfert des électrons à travers cette région. Ainsi, il vient

$$I_{cp} = \beta_1\gamma_1 I_1, \tag{12.34}$$

$$I_{cn} = \beta_2\gamma_2 I_2. \tag{12.35}$$

Le courant de collecteur total a pour valeur

$$I_c = I_{cs} + I_{cp} + I_{cn} = I_{cs} + \beta_1\gamma_1 I_1 + \beta_2\gamma_2 I_2. \tag{12.36}$$

Un dynistor est dans un état stable lorsque les courants totaux à travers toutes les trois jonctions *p-n* sont égaux l'un à l'autre:

$$I_1 = I_c = I_2 = I. \tag{12.37}$$

Dans ce cas

$$I = I_{cs} + (\beta_1\gamma_1 + \beta_2\gamma_2)\, I. \tag{12.38}$$

En résolvant cette équation par rapport à I, on obtient

$$I = \frac{I_{cs}}{1-(\beta_1\gamma_1+\beta_2\gamma_2)}. \tag{12.39}$$

Le dynistor est réalisé de telle sorte qu'à l'état bloqué le paramètre $(\beta_1\gamma_1 + \beta_2\gamma_2)$ soit inférieur à l'unité. Cependant à mesure que la tension sur le dynistor et donc sur la jonction collectrice augmente, les coefficients de transfert β_1 et β_2 s'accroissent parce que la région de charge d'espace de cette jonction s'élargit (fig. 12.22, *b*) et les distances que doivent parcourir les porteurs minoritaires injectés à travers les jonctions émettrices diminuent. La croissance de β_1 et β_2 conduit, comme il s'ensuit de la relation (12.39), à une augmentation du courant I à travers le dynistor. En outre, le courant dans le dynistor peut s'accroître par suite d'une ionisation par choc dans la jonction collectrice qui intervient dans le cas de fortes polarisations négatives. (La formule (12.36) ne tient pas compte de cet effet.)

La croissance du courant I s'accompagne aussi de la croissance des coefficients d'injection γ_1 et γ_2 (v. § 12.2). Ainsi, le paramètre $(\beta_1\gamma_1 + \beta_2\gamma_2)$ augmente en même temps que la tension appliquée au dynistor. Pour une certaine valeur de la tension appliquée au dynistor ce paramètre doit devenir égal à l'unité et alors le courant passant par l'appareil doit, suivant (12.39), tendre vers l'infini.

Or, même avant cet instant, le régime de fonctionnement du dynistor devient instable et le processus de sa commutation commence. En effet, quand $(\beta_1\gamma_1 + \beta_2\gamma_2) \to 1$, la somme des courants d'élec-

trons et de trous rassemblés par la jonction collectrice commence à dépasser le courant $I_{\acute{e}}$ passant par les jonctions émettrices:

$$I_c = I_{cs} + (\beta_1\gamma_1 + \beta_2\gamma_2)\, I_{\acute{e}} > I_{\acute{e}}. \qquad (12.40)$$

En d'autres termes, le nombre de trous transférés maintenant par la jonction collectrice dans la région p_2 est supérieur à celui qui est nécessaire pour la recombinaison avec des électrons injectés dans cette région et un nombre excédentaire d'électrons est transféré dans la région n_1.

Cela signifie que des porteurs de signes contraires, les électrons et les trous non compensés, commencent à s'accumuler près de la jonction n_1-p_2. Les trous en excès arrivant vers la frontière entre la région de charge d'espace de la jonction collectrice et la région p_2 et les électrons en excès de l'autre côté de la jonction n_1-p_2 réduisent la région de charge d'espace de cette jonction. Dans ces conditions, la hauteur de la barrière de potentiel de la jonction collectrice diminue de façon continue et donc la chute de tension inverse dans cette jonction diminue tandis que la polarisation directe des jonctions émettrices augmente. Si la diminution de β due à la diminution de la tension sur la jonction n_1-p_2 n'est pas compensée par l'augmentation de γ par suite de la croissance de la polarisation directe des jonctions émettrices, le paramètre $(\beta_1\gamma_1 + \beta_2\gamma_2)$ diminue et l'appareil reprend son état stationnaire. Cet état correspond pourtant à un courant fort pour une chute de tension dans la structure diminuée par suite de la diminution de la polarisation inverse de la jonction collectrice. (Une certaine croissance de la polarisation directe des jonctions émettrices n'a pas d'effet.)

Si le courant s'accroît encore, il se produit dans les régions de base du dynistor une accumulation complémentaire des porteurs majoritaires dont la charge n'est pas compensée par la charge des porteurs minoritaires et donc la tension de polarisation inverse de la jonction collectrice diminue encore. C'est ainsi que sur la courbe caractéristique courant-tension du dynistor il apparaît une partie à résistance négative. A une certaine étape (souvent même lorsque la caractéristique n'a pas encore atteint sa partie à résistance négative), le processus d'accumulation des porteurs majoritaires dans les bases de l'appareil commence à se développer d'une façon incontrôlable parce que le paramètre $(\beta_1\gamma + \beta_2\gamma_2)$ reste tout le temps voisin de l'unité ou même la dépasse. Le processus transitoire continue tant que la hauteur de la barrière de potentiel de la jonction collectrice ne devienne proche de sa valeur d'équilibre et ensuite inférieure à cette dernière. (Autrement dit, cette jonction se trouvera polarisée en sens direct.) Dans ce cas, la jonction collectrice n'est plus en mesure de retenir les trous dans la région p_2 et les électrons dans la région n_1 de sorte que ces porteurs commencent à remplir les régions n_1 et p_2 respectivement.

Finalement, au lieu de la répartition des porteurs minoritaires représentée pour l'état bloqué du dynistor (sur la fig. 12.22, *b*), il apparaît dans ces régions une répartition caractéristique de son état conducteur (fig. 12.22, *c*). Dans cet état toutes les trois jonctions *p-n* du dynistor sont polarisées en sens direct. Les régions n_1 et p_2 se trouvent fortement « inondées » par des porteurs de charge minoritaires, ce qui provoque une diminution des coefficients d'injection γ des jonctions émettrices. Le courant à travers les régions n_1 et p_2 est transporté maintenant non seulement par diffusion des porteurs minoritaires dans ces régions mais aussi par dérive des porteurs majoritaires et minoritaires dont les concentrations dans ces régions ne diffèrent que peu.

L'état conducteur du dynistor est stable parce que le paramètre $(\beta_1\gamma_1 + \beta_2\gamma_2)$ est maintenant inférieur à l'unité: les coefficients γ sont diminués et l'action séparatrice de la jonction collectrice est devenue moins efficace. La résistance de l'ensemble du dynistor à l'état conducteur est faible (courbe *2* de la figure 12.21) et pour éviter le passage des courants trop intenses à travers l'appareil, il faut monter en série avec lui une résistance de limitation de courant. Le courant du dynistor à l'état conducteur se détermine par la tension de la source de polarisation et la résistance du circuit comme l'indique la figure 12.23. Si l'on diminue le courant traversant un dynistor conducteur, le nombre de porteurs minoritaires injectés à travers les émetteurs diminuera lui aussi. Lorsque pour un certain courant I_{bl} leur nombre deviendra insuffisant pour maintenir la polarisation directe de la jonction collectrice, la résistance de cette jonction augmente et elle sera de nouveau polarisée en sens inverse. Dans ce cas la tension de polarisation appliquée au dynistor est redistribuée: seule une petite partie de cette tension chutera dans les jonctions émettrices et l'appareil passera à l'état bloqué.

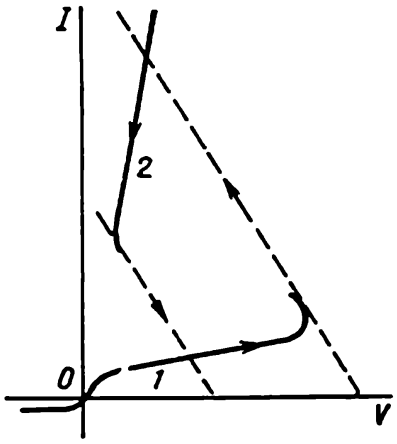

Fig. 12.23. Mise en et hors circuit d'un dynistor.

Les processus de bloquage et de débloquage du dynistor sont non stationnaires. Ils se développent en avalanche de façon non contrôlée. C'est pourquoi la caractéristique courant-tension du dynistor se compose de deux parties (*1* et *2* sur les figures 12.21 et 12.23) correspondant à ses états stationnaires, stables. Pour ce qui est des courants et des tensions du dynistor lors de sa commutation, ils dépendent de la résistance du circuit, de la tension de la source de polarisation, des capacités et des inductances du circuit.

Dans le cas de la polarisation inverse (le pôle négatif de la source étant relié à la région p_1), le courant du dynistor est déterminé par la résistance des jonctions émettrices. Il est faible jusqu'à la tension de claquage de ces jonctions (courbe *3* de la figure 12.21).

Les dynistors sont employés en radiotechnique et en automatique comme éléments de commutation. C'est un relais électronique particulier qui fonctionne lorsque la tension appliquée atteint la valeur de branchement V_{br}.

La tension de branchement du thyristor peut être commandée à l'aide d'une troisième sortie, supplémentaire — de l'électrode de commande (borne « Com » sur la figure 12.24, *a*). La polarité de la tension de polarisation appliquée à l'électrode de commande doit

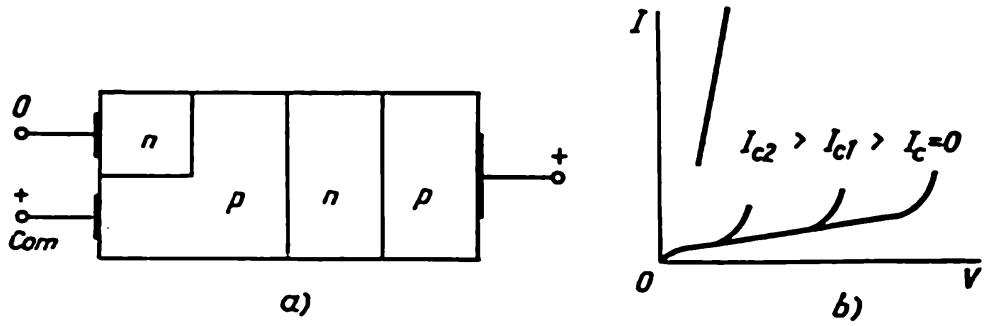

Fig. 12.24. Constitution (*a*) et caractéristiques courant-tension (*b*) d'un thyristor commandé.

être telle que le courant d'injection de la jonction *p-n* émettrice la plus proche augmente et donc le courant du thyristor tout entier augmente lui aussi. La figure 12.24, *b* montre une famille de caractéristiques courant-tension d'un thyristor-triode pour différents courants I_{com} à travers l'électrode de commande.

Le courant à travers un thyristor bloqué peut être augmenté et le thyristor peut être rendu conducteur si l'on éclaire ses régions de base. Un tel appareil est appelé *photothyristor*.

Ces dernières années, dans l'électronique intégrée on utilise comme éléments de commutation des films minces (près de 1 μm) de semiconducteurs vitrifiés, par exemple $Fe_{50}As_{30}Si_{10}Ge_{10}$. Leur conductivité électrique aux températures ordinaires est de 10^{-5} à 10^{-8} $ohm^{-1} \cdot cm^{-1}$ et à l'état conducteur elle est de plusieurs ordres de grandeur plus élevée. Une des explications d'un tel effet est la formation par fusion des canaux entre les électrodes et une modification de leur structure cristalline qui conduit à la formation de « ponts » conducteurs dans le verre. Cependant il n'existe pas jusqu'à présent de théorie complète de ce phénomène, bien que dans le monde entier on attache une grande importance à de tels commutateurs.

§ 12.7. Dispositifs à semiconducteurs comme éléments des circuits intégrés

Les matières concernant l'électronique intégrée sont étudiées dans des cours spéciaux. Mais il nous a semblé raisonnable de donner dans ce cours d'introduction à l'électronique des solides des renseignements élémentaires sur les circuits intégrés et l'utilisation des dispositifs à semiconducteurs comme éléments de ces circuits.

Le développement de la science et de la technique a exigé d'élaborer des appareils radiotechniques complexes, hautement efficaces et fiables, contenant des centaines de milliers et parfois même des dizaines de millions d'éléments réunis en un circuit commun unique. La recherche des voies permettant de résoudre ce problème a conduit à la mise au point des éléments et organes électroniques *miniaturisés* sur la base du corps solide et à la *microminiaturisation* des appareillages radioélectroniques qui a donné naissance à une nouvelle branche de l'électronique — la *microélectronique* dont l'objet principal est précisément la réalisation des circuits et dispositifs électroniques *à haute fiabilité*, *microminiatures*, *économiques et à bas prix de revient*. Pour résoudre ce problème, on utilise non seulement les méthodes d'une simple réduction des dimensions des composants radio-électriques distincts avec la conservation des procédés traditionnels d'assemblage d'organes fonctionnels correspondants à partir de ces composants, mais également des principes qualitativement nouveaux de construction et de fabrication des circuits électroniques formant un tout lorsque les éléments actifs, les éléments passifs et les éléments de connexion sont formés dans des microvolumes d'un cristal semiconducteur ou sur sa surface oxydée, au cours d'un processus technologique unique. Le circuit électronique ainsi réalisé, destiné à l'exécution d'une fonction déterminée, est appelé *circuit intégré* (CI). Ses éléments constitutifs (transistors, diodes, condensateurs, résistances, etc.) n'ont pas de sorties extérieures et ne peuvent pas être considérés comme des articles indépendants. Le circuit dans son ensemble comporte un nombre nécessaire de plots de soudure servant à sa connexion avec les sorties du boîtier qui le protège contre les dégradations mécaniques et l'influence de l'environnement. Les éléments du circuit sont reliés entre eux soit à l'intérieur du cristal, au moyen de couches semiconductrices spécialement réalisées à cet effet, soit à l'aide de pistes métalliques portées sur la surface oxydée du cristal d'ordinaire par les méthodes d'évaporation sous vide.

Suivant leur destination fonctionnelle les circuits intégrés se divisent en *circuits logiques* (*numériques*) et *circuits linéaires* (*analogiques*). Les circuits intégrés logiques fonctionnent en régime de commutation et sont employés dans les calculateurs électroniques, les systèmes d'automatisme et dans les dispositifs de traitement discret de l'information. Les circuits intégrés linéaires fonctionnent en

régimes linéaires des amplificateurs et des oscillateurs ou effectuent des transformations non linéaires des signaux d'entrée dans les montages mélangeurs, détecteurs, etc. De nombreux schémas des amplificateurs et des convertisseurs sont à base d'amplificateur opérationnel.

Du point de vue de l'intégration des circuits, c'est-à-dire de la réunion d'un grand nombre d'éléments, on convient de caractériser les circuits intégrés par les paramètres suivants : la *densité de composants* et le *degré* ou le *niveau d'intégration*.

La densité de composants se détermine par le nombre d'éléments par unité de volume (1 cm^3) du CI. Il est naturel que la croissance de la densité de composants est liée à la réduction de dimensions des éléments de CI. Cela est particulièrement important pour les circuits numériques équipant les calculateurs électroniques parce que la diminution des dimensions des éléments de CI augmente leur rapidité de fonctionnement et diminue la puissance dissipée. Dans les montages amplificateurs, la diminution de la longueur des connexions entre les éléments a pour effet de réduire les inductances parasites des sorties et la faible dimension des éléments diminue les capacités parasites, ce qui permet d'élever la fréquence de travail pour de petites puissances dissipées.

Le degré d'intégration se détermine par le nombre d'éléments contenus dans le CI. On a adopté la gradation suivante des CI d'après le degré d'intégration (tableau 12.1).

Tableau 12.1

Degré d'intégration	Nombre d'éléments dans le CI	Appellation du circuit
Premier	Jusqu'à 10	CI à faible niveau d'intégration (SSI) *)
Deuxième	De 10 à 100	CI à moyenne échelle (MSI) **)
Troisième	De 100 à 1000	CI à grande échelle (LSI) ***)
Quatrième	Plus de 1000	CI à très haut niveau d'intégration (VLSI) ****)
Cinquième	Plus de 10 000	

*) De l'anglais *single scale integration*.
**) De l'anglais *medium scale integration*.
***) De l'anglais *large scale integration*.
****) De l'anglais *very large scale integration*.

L'augmentation du degré d'intégration des CI permet d'améliorer la qualité et la fiabilité du matériel radio-électronique et d'abaisser son prix de revient. L'amélioration de la fiabilité est obtenue grâce à une forte diminution du nombre de connexions extérieures qui sont des endroits les plus vulnérables et les moins sûrs des cir-

cuits. La baisse du prix de revient est due à une diminution importante du nombre de boîtiers dans lesquels étaient rendus étanches les composants discrets des circuits ordinaires, à la diminution du nombre d'opérations d'assemblage et des connexions internes ainsi qu'à l'utilisation de la méthode complexe de fabrication de divers éléments du circuit au cours d'un cycle technologique unique.

Suivant le procédé de fabrication et la structure obtenue on distingue des circuits intégrés *monolithiques* et les circuits intégrés *hybrides*.

Les *CI monolithiques* sont réalisés dans une plaquette de semiconducteur monocristallin (le plus souvent, de silicium) par les méthodes d'implantation locale des impuretés convenables pour créer des éléments actifs (transistors et diodes) et passifs (résistances, condensateurs), pour les isoler l'un de l'autre et pour obtenir des connexions entre les éléments. La dimension d'un cristal de CI distinct, qu'on appelle pastille intégrée, peut varier de quelques dixièmes de millimètre à quelques millimètres, alors que les dimensions des éléments constitutifs de CI sont généralement comprises entre quelques microns et quelques dizaines de microns et, dans certains circuits, elles sont réduites à quelques dixièmes de micron.

Dans les *CI hybrides* les éléments passifs sont fabriqués sur un support diélectrique par les méthodes de la technologie à couches tandis que les éléments actifs — les diodes, les transistors et les CI monolithiques sans boîtier sont fixés sur le support et connectés aux éléments passifs et entre eux-mêmes. Dans les circuits hybrides à couches épaisses les éléments passifs et les connexions sont obtenus en déposant sur les supports, à travers des pochoirs, des pâtes spéciales, conductrices, résistives et isolantes et en les soumettant à la cuisson. Dans les circuits hybrides à couches minces, les films correspondants sont déposés par évaporation sous vide, par pulvérisation cathodique ou par une autre méthode quelconque. La réunion des éléments actifs en un circuit se fait par soudage, par thermocompression et par d'autres procédés.

Notions sur la technologie planar. Les CI monolithiques sont fabriqués d'ordinaire selon une *technologie* dite *planar* qui comporte les opérations principales suivantes:

1. L'*oxydation de la surface* de la plaquette de silicium qui est effectuée dans des fours spéciaux en atmosphère d'oxygène sec ou humide à des températures de 900 à 1200 °C. Il en résulte la formation, sur la surface de silicium, d'un film de dioxyde de silicium SiO_2 qui présente des propriétés importantes au point de vue de son utilisation pour la fabrication des CI. Premièrement, le film de SiO_2 est un parfait diélectrique et de ce fait peut être utilisé pour l'isolement de la métallisation portée sur la surface de l'oxyde, des éléments de CI formés dans la masse ou sur la surface du cristal. Deuxièmement, ce film est pratiquement imperméable à de nombreuses impuretés

dopantes utilisées pour la réalisation des CI. Le SiO_2 peut donc être utilisé comme *masque* pour la protection des régions nécessaires du cristal contre la pénétration des impuretés de dopage lors de l'exécution du processus de diffusion. Suivant la température à laquelle sera implantée l'impureté, sa nature et la durée du processus de diffusion, l'épaisseur du film de SiO_2 varie de 100 à 1000 nm.

2. Pour obtenir le « dessin » d'un CI futur, c'est-à-dire pour obtenir les représentations des éléments actifs et passifs du circuit, des plots de soudure, des pistes de connexion, etc., on a recours à la *photolithographie* qui rappelle dans une certaine mesure le processus de photographie ordinaire. Sur la surface oxydée de la plaquette de silicium on dépose une couche mince (épaisse de quelques centaines de nanomètres) d'un vernis spécial appelé *photorésist*, sensible à la lumière ultraviolette. Après le séchage du photorésist on lui applique fortement, dans une installation spéciale, un *cliché photographique* en verre qui porte l'image négative du dessin à transporter sur la plaquette de silicium. A travers ce cliché les parties non couvertes du photorésist sont éclairées par la lumière ultraviolette et éliminées ensuite à l'aide d'un décapant spécial, en découvrant dans des « fenêtres » ainsi formées une couche d'oxyde. Au moyen d'un autre décapant qui dissout le film d'oxyde mais n'agit pas sur le photorésist, on élimine dans les fenêtres ouvertes le film d'oxyde et on met à nu la surface du silicium. Après l'enlèvement du photorésist restant, le dessin qui était réalisé sur le cliché photographique est transporté sur le film d'oxyde sous la forme de fenêtres ouvertes et de régions fermées (masquées) par des îlots de couche d'oxyde.

3. A travers les fenêtres ouvertes, on effectue une *diffusion localisée des impuretés dopantes* pour conférer aux couches de silicium situées sous ces fenêtres des propriétés électrophysiques désirées (conductibilité de type voulu, concentration de porteurs et autres).

Isolement des éléments des circuits intégrés. Comme tous les éléments d'un circuit intégré monolithique sont réalisés sur une même plaquette — le substrat, ils doivent être bien isolés les uns des autres pour prévenir une influence parasite de l'un sur l'autre. Cet isolement s'obtient soit à l'aide de jonctions *p-n* polarisées en inverse, soit à l'aide de couches diélectriques.

La figure 12.25 montre le schéma de la formation des îlots de silicium *n* isolés à l'aide de jonctions *p-n* de séparation.

Le matériau de départ est une plaquette de silicium *p* portant une couche *n* épitaxiale. A leur surface frontière il se forme une jonction *p-n*. Pour obtenir des îlots de silicium *n* distincts, isolés l'un de l'autre, la surface de la couche épitaxiale est oxydée et on y ouvre, par le procédé de photolithographie décrit plus haut, des fenêtres comme l'indique la figure 12.25, *a*. A travers ces fenêtres, on effectue la diffusion d'une impureté acceptrice (par exemple du bore) sur toute l'épaisseur de la couche épitaxiale. Après l'élimination du

film d'oxyde de masquage il reste dans la plaquette des îlots de film épitaxial de silicium *n* séparés du silicium *p* par la jonction *p-n* isolante (fig. 12.25, *b*). C'est dans de tels îlots qu'on forme des éléments distincts du circuit intégré.

La figure 12.26 illustre la technologie de l'*isolement diélectrique* des éléments d'un circuit intégré à réaliser. Sur la surface de silicium *n* on forme une couche d'oxyde (fig. 12.26, *a*), on y ouvre, par les méthodes de la photolithographie, des fenêtres à travers lesquelles on effectue un décapage profond (fig. 12.26, *b*). Après une deuxième oxydation (fig. 12.26, *c*) on dépose une couche de silicium polycristallin épaisse de 100 à 200 μm (fig. 12.26, *d*). Le silicium monocristallin est ensuite éliminé jusqu'à la couche d'oxyde. Il se forme finalement des îlots de silicium *n* isolés les uns des autres par des couches diélectriques de SiO_2.

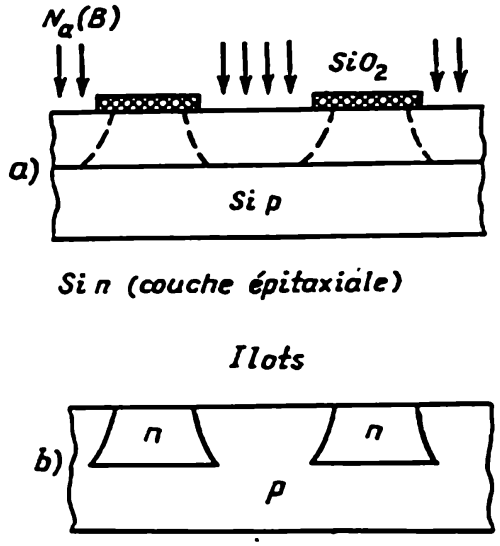

Fig. 12.25. Schéma de la formation des îlots isolés de silicium de type *n* à l'aide de jonctions *p-n* de séparation.

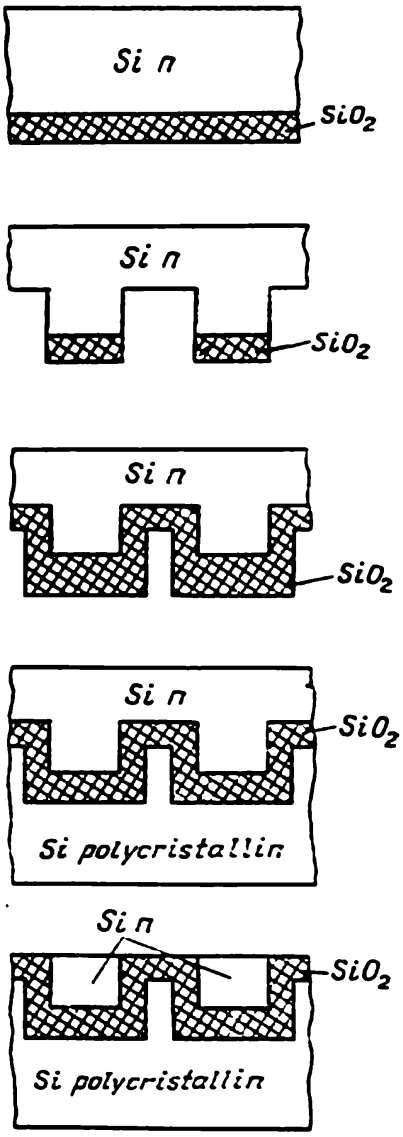

Fig. 12.26. Schéma de la formation des îlots de silicium de type *n* isolés l'un de l'autre par des couches diélectriques (SiO_2).

Eléments actifs et passifs des circuits intégrés monolithiques. Le *transistor bipolaire* est l'élément actif le plus répandu des circuits intégrés monolithiques. On utilise le plus souvent les transistors

de type *n-p-n* dont la mobilité des porteurs minoritaires (des électrons) dans la base est plus élevée que dans les transistors de type *p-n-p*.

Examinons sommairement le schéma de la formation d'un tel transistor sur l'exemple des circuits intégrés au silicium à isolement par jonctions *p-n* (fig. 12.27).

Le matériau de départ est une plaquette de silicium *p* à résistivité de l'ordre de 10 ohms·cm. La plaquette est oxydée et, aux endroits où seront créés les collecteurs, on ouvre par la méthode de la photolithographie, des fenêtres à travers lesquelles on effectue l'implantation

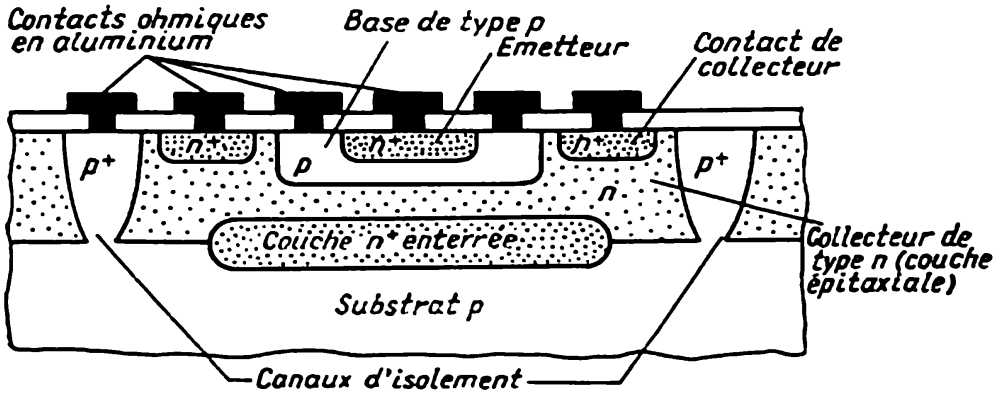

Fig. 12.27. Structure d'un transistor bipolaire d'un circuit intégré au silicium à isolement par jonction *p-n*.

d'une impureté donatrice (généralement de l'arsenic) pour la formation des couches enterrées n^+. Ces couches sont destinées à *réduire la résistance du collecteur*.

Après l'introduction de l'impureté donatrice et la formation des couches enterrées n^+ le film d'oxyde est éliminé par décapage et sur la surface propre du silicium *p* (substrat) on dépose une couche épitaxiale *n* qui remplira le *rôle de la région collectrice* pour tous les transistors à former sur la plaquette.

Ensuite, on oxyde la couche épitaxiale et on y ouvre par la méthode de la photolithographie des fenêtres à travers lesquelles on forme par diffusion localisée de l'impureté acceptrice des *canaux d'isolement*. L'impureté doit diffuser dans toute la profondeur de la couche épitaxiale jusqu'à la surface du substrat. Il en résulte que toutes les régions *n* sont séparées l'une de l'autre et du substrat par des jonctions *p-n* d'isolement.

Le stade suivant du processus de fabrication est la formation des *régions p de base*. A cet effet, le film d'oxyde déposé antérieurement est éliminé et la plaquette est oxydée de nouveau. Dans le film d'oxyde obtenu on ouvre des fenêtres pour les régions de base à travers lesquelles on effectue une diffusion localisée de l'impureté acceptrice. Pour obtenir la résistance désirée de ces régions, il est nécessai-

re de contrôler avec soin et précision la concentration en surface de l'impureté introduite et la profondeur de sa pénétration.

Après la formation des régions de base la couche d'oxyde est éliminée et la plaquette est oxydée de nouveau. Dans la couche d'oxyde nouvellement obtenue on ouvre, par la méthode de la photolithographie, des fenêtres pour les *régions émettrices* et les *contacts de collecteur*. A travers ces fenêtres on introduit l'impureté donatrice pour la formation des régions n^+ des émetteurs et des contacts du collecteur.

Le stade final de fabrication d'un transistor est la création des plots de soudure. A cet effet, la couche d'oxyde est éliminée et toute la plaquette est oxydée de nouveau. Dans la couche d'oxyde obtenue on ouvre, par la méthode de la photolithographie, des fenêtres pour les contacts du transistor. Ordinairement, en employant l'évaporation sous vide de l'aluminium, on forme des contacts et des plots de soudure vers les régions d'émetteurs, de bases et de collecteurs du transistor. Pour éliminer la métallisation en aluminium aux surfaces de la plaquette où cette métallisation n'est pas nécessaire, on effectue encore une photolithographie.

Diodes des CI. Elles présentent d'ordinaire une structure de transistor permettant de faire varier les modes de leur branchement suivant les exigences imposées aux paramètres (rapidité de fonctionnement, tension de claquage, intensité de courant direct, etc.).

Résistances diffusées. Dans les CI monolithiques ce sont généralement les régions de base qu'on utilise comme résistances. Un inconvénient de telles résistances est une gamme limitée des valeurs ohmiques possibles et l'impossibilité de satisfaire à des tolérances sévères sur ces valeurs, les fortes valeurs des coefficients de température de résistance, la possibilité de formation d'éléments parasites, une limitation imposée à la tension appliquée.

Condensateurs diffusés. Ce sont les capacités des jonctions *p-n* des structures transistors (d'émetteur ou de collecteur) qu'on utilise comme condensateurs dans les CI monolithiques. Le facteur de qualité (l'inverse de la tangente de pertes diélectriques) de tels condensateurs est en règle générale beaucoup plus faible que celui des condensateurs discrets à film diélectrique.

CI MOS. Ils sont réalisés sur la base de transistors à effet de champ (v. fig. 12.15) à grille isolée. Les transistors sont utilisés comme éléments actifs et comme résistances de charge, ce qui assure de belles qualités technologiques de fabrication de tels CI. Dans les CI MOS de même type de conductibilité du canal l'isolement des éléments l'un de l'autre n'est pas exigé parce que leurs jonctions *p-n* sont toujours polarisées en inverse. Les dimensions des transistors à effet de champ sont inférieures d'un ordre de grandeur environ à celles des transistors bipolaires. Cela permet d'obtenir une densité de composants plus élevée. C'est pourquoi les structures MOS sont large-

ment employées dans la construction des circuits de très grande intégration (VLSI).

Les circuits intégrés, aussi bien à transistors bipolaires qu'à transistors MOS, sont fabriqués suivant la technologie planar simultanément en un grand nombre d'exemplaires. Sur une même plaquette de 100 mm de diamètre et plus on forme à la fois jusqu'à quelques milliers de CI. Après leur réalisation au cours d'un cycle technologique unique les CI contenus dans la plaquette sont séparés les uns des autres en cristaux distincts (que l'on appelle aussi *chip*, *pastille* ou *puce*) dont chacun représente un CI distinct. Après la vérification de leur bon fonctionnement et la mise au rebut des pièces défectueuses les CI sont mis en boîtier.

CI hybrides. Comme nous l'avons vu, les éléments passifs des CI monolithiques ont certains inconvénients sérieux qu'on peut affaiblir dans une mesure importante si l'on utilise pour leur fabrication une technologie à couches épaisses ou minces.

Résistances à films. Elles sont usinées sous forme de bandes de configuration donnée par la méthode de l'évaporation sous vide des métaux purs (chrome, tantale et autres), des alliages métalliques (nichrome et autres) et des matériaux métallocéramiques (cermets). Les couches épaisses sont réalisées par la méthode de sérigraphie. Les résistances à couches épaisses présentent par rapport aux résistances diffusées toute une série d'avantages substantiels: une faible valeur du coefficient de température de résistance, une large gamme de valeurs nominales et la possibilité de réglage précis à la valeur nominale, l'absence d'éléments parasites, une haute stabilité avec le temps.

Condensateurs à couches. Ils sont fabriqués en déposant sur le substrat une couche métallique mince, puis une couche diélectrique mince et ensuite de nouveau une couche métallique mince. Pour augmenter la capacité, on utilise souvent des structures multicouches avec montage en parallèle des condensateurs ainsi obtenus. Par leurs paramètres, les condensateurs à couches sont en règle générale supérieurs de façon notable aux condensateurs diffusés.

Technologie combinée. Les paramètres nettement plus hauts des éléments passifs obtenus par les méthodes de la technologie à couches ont obligé de passer dans certains cas à la fabrication des CI par la technologie combinée dans laquelle les éléments actifs de CI sont formés dans la masse du cristal semiconducteur et les éléments passifs sur la surface du cristal (protégé le plus souvent par un film d'oxyde) par les méthodes de la technologie à couches. La technologie combinée est utilisée particulièrement souvent pour la fabrication des CI analogiques dans lesquels les paramètres doivent faire face à des exigences sévères.

Technologie hybride. Le développement parallèle de la technologie à couches et de la technologie à semiconducteurs a conduit au

développement d'une *technologie hybride* et à la création d'une large classe de circuits intégrés hybrides (CIH). Les avantages principaux que ces circuits présentent par rapport aux circuits monolithiques sont les suivants:

a) la possibilité de réaliser une large classe de CI numériques et analogiques pour un cycle relativement court de leur élaboration;

b) le caractère universel de la méthode de conception et de construction des microcircuits assurant la possibilité d'utiliser des composants actifs qui diffèrent par la destination fonctionnelle et l'exécution constructive (CI sans boîtier, transistors MOS, arrangements de diodes et de transistors et autres);

c) la possibilité d'obtenir des éléments passifs à couches d'une large nomenclature avec des tolérances sévères;

d) une haute efficacité économique de fabrication de petits lots de circuits;

e) un pourcentage relativement haut de sortie des circuits intégrés bons grâce à la possibilité de rejet des éléments défectueux avant le montage.

Pour toutes ces raisons, les circuits intégrés hybrides trouvent de nombreuses applications dans la construction des appareillages microélectroniques.

ANNEXES

1. Constantes physiques

Constante	Notation	Valeur
Constante de Boltzmann	k	$1{,}38\cdot10^{-23}$ J/K $=8{,}63\cdot10^{-5}$ eV/K
Constante de Planck	h	$6{,}63\cdot10^{-34}$ J·s
	$\hbar$	$1{,}055\cdot10^{-34}$ J·s
Permittivité du vide	ε_0	$8{,}85\cdot10^{-12}$ F/m
Perméabilité du vide	μ_0	$1{,}26\cdot10^{-6}$ H/m
Vitesse de la lumière dans le vide	c	$3\cdot10^{8}$ m/s
Charge de l'électron	e	$1{,}6\cdot10^{-19}$ C
Masse de l'électron	m_e	$9{,}11\cdot10^{-31}$ kg
Magnéton de Bohr	μ_B	$9{,}27\cdot10^{-24}$ A·m²
Magnéton nucléaire	μ_{nuc}	$5{,}05\cdot10^{-27}$ A·m²

2. Propriétés du germanium et du silicium

Grandeur physique	Germanium	Silicium
Numéro atomique	32	14
Masse molaire, g	72,59	28,08
Température de fusion, °C	936	1420
Constante du réseau cristallin, nm	0,565	0,543
Concentration des atomes, cm^{-3}	$4{,}4\cdot10^{22}$	$5\cdot10^{22}$
Masse volumique à 300 K, g/cm³	5,327	2,329
Largeur de la bande interdite à 300 K, eV	0,67	1,1
Mobilité des électrons dans un semiconducteur intrinsèque à 300 K, cm²/(V·s)	3800	1350
Idem pour les trous	1800	480
Coefficient de diffusion des électrons dans un semiconducteur intrinsèque à 300 K, cm²/s	93	35
Idem pour les trous	44	12,5
Masse effective de la densité des états dans la bande de conduction	$0{,}55m_e$	$1{,}1m_e$
Idem dans la bande de valence	$0{,}39m_e$	$0{,}59m_e$
Concentration des porteurs intrinsèques à 300 K, cm^{-3}	$2{,}4\cdot10^{13}$	$1{,}5\cdot10^{10}$
Résistivité d'un semiconducteur intrinsèque à 300 K, ohm·cm	50	$2{,}5\cdot10^{5}$
Permittivité relative	16	12
Indice de réfraction	4	3,5

BIBLIOGRAPHIE

Бонч-Бруевич В. Л., Калашников С. Г. Физика полупроводников. М., 1977.
Буккель В. Сверхпроводимость. М., 1975.
Зи С. Физика полупроводниковых приборов. М., 1984.
Киттель Ч. Введение в физику твердого тела. М., 1978.
Панков Ж. Оптические процессы в полупроводниках. М., 1973.
Пасынков В. В., Чиркин Л. К. Полупроводниковые приборы. М., 1986.
Smith R. Semiconductors, Cambridge, 1959.

INDEX

Absorption de la lumière
— —, coefficient d' 209
— — due aux impuretés 216
— — par des porteurs de charges libres 214
— — propre 21
— —, section différentielle 217
Action de masse, loi d' 115
Aimantation 157
— spontanée 168
Anisotropie cristallographique magnétique 171
Antiferromagnétisme 172
Appauvrissement, région d' 329
Approximation harmonique 48
— monoélectronique 74
Atome disloqué 33
— en position interstitielle 97
Axe de difficile aimantation 172
— de facile aimantation 172

Bande de conduction 87
— d'énergie 76
— d'impureté 116
— interdite 87
— —, largeur optique 88, 212
— de valence 87
Bardeen-Cooper-Schrieffer, théorie de 139
Barkhausen, effet 178
Bloch, théorème de 77
Bohr, magnéton de 163
Boltzmann, équation cinétique de 128
—, répartition de 106
Born, énergie de répartition suivant 24
Bragg-Wulff, loi de 15, 20
Bravais, réseau de 7
Brillouin, fonction de 165
—, théorie du paramagnétisme de 164
—, zone de 49, 78
Burgers, contour de 38
—, vecteur de 37, 38, 40
Burstein, déplacement de 213

Capacité de barrière 280
— de diffusion 287
Capture, section efficace de 197
Caractéristique courant-tension d'une diode inversée 296
— — — — Schottky 255
— — — — tunnel 297
— — — d'un dynistor 337
— — — d'une jonction *p-n* 284
— — — d'une photodiode 301
— — — en S 336
— — — de la structure MDM 258, 261
— — — d'un thyristor commandé 342
— — — d'un transistor bipolaire 316, 319
— — — — MOS 333, 336
Centres de couleur ou de coloration 35
Chaleur massique du gaz électronique 120
— — du réseau cristallin 61
Champ critique dans un supraconducteur 143
— de départ 179
Child-Langmuir, équation de 245
Circuit intégré 343
— — hybride 345, 350
— — linéaire 343
— — logique 343
— — monolithique 345
— — MOS 349
Cisaillement, formation de 41
Claquage d'une jonction collectrice 319
— — *p-n* 292
— — — par avalanche 295
— — — par effet tunnel 296
— — — thermique 293
Coefficient de conductibilité thermique 68
Collapsus, champ de 189
Concentration d'équilibre de défauts dans le cristal 33

Concentration de phonons 61
— de porteurs de charges dans les semiconducteurs extrinsèques 11, 112
— — — — intrinsèques 107
Condensateur à couches 350
— diffusé 349
Conditions aux limites périodiques 50
Conductance différentielle négative 231
Conductibilité d'une impureté 35
— intrinsèque, température de passage à la 114
— des semiconducteurs 135
Conduction d'obscurité 219
Conductivité électrique des diélectriques 86
— — des métaux 86, 129, 136
Constantes d'anisotropie magnétique 172
Contact de bloquage 250
— de deux métaux 246
— entre un métal et un semiconducteur 249
— neutre 266
— ohmique 256
Continuité, équation de 201
Cooper, paires de 140
Couche appauvrie 250
— enrichie 252, 330
Courant direct d'une diode Schottky 255
— limité par la charge d'espace 257
— de saturation 255, 284
— de seuil d'une diode laser 310
— à travers la jonction *p-n* 282
Courbe de dispersion 49
— — d'un électron dans le cristal 78, 80
— — — libre 80
— — pour les vibrations du réseau cristallin 52
Cristal 7
— ionique 24
— métallique 28
— méthode de la détermination de la structure du 20
— moléculaire 29
— parfait 75
Curie, constante de 159
—, loi de 159
—, point ferromagnétique de 160
—, température antiferromagnétique de 172
Curie-Weiss, loi de 161
Cycle d'hystérésis 160
Cyclotron, fréquence 98
Debye, longueur de 330
—, température caractéristique de 61
—, théorie de la chaleur massique du réseau cristallin 60
Défauts d'assemblage 45
— bidimensionnels 44
— de Frenkel 32, 33
— linéaires 36
— de Schottky 33
— tridimensionnels 46
— du type soustraction 46
Déformation des cristaux 41
Densité de composants 344
— de courant critique dans un supraconducteur 143
— de répartition des modes suivant les fréquences 55
Déplétion de l'impureté, température de 112
Diamagnétisme 161
— des supraconducteurs 139
Différence de potentiel de contact 246
Diffraction, conditions de 18
— dans un cristal 16
Diffusion, coefficient effectif de 203
— conjointe des porteurs de charges hors d'équilibre 202
— des électrons sur les atomes d'impureté 126
— — sur les vibrations thermiques du cristal 126
— localisée des impuretés dopantes 346
—, longueur de 205
— phonon-phonon 68
Dilatation thermique 63
—, coefficient de température de 66
Diode à accumulation de charges 289
— à avalanche 298
— à base longue 336
— luminescente 305
— —, rendement quantique interne 305
— semiconductrice 277
— trijonction 337
— tunnel 297
Dislocations, axe ou ligne de 38
— coins 36
—, densité de 41
—, modèle de galets de 43
— partielles 38
— totales 38
— vis 38
Dispersion, relation de 49
Distance réticulaire 15

Domaines électrostatiques 235
— magnétiques 168, 175
— — cylindriques 187
—, paroi de 176
—, processus de déplacement des frontières des 177
Donneurs 95
Dulong et Petit, loi de 59
Durée de vie 193
Dynistor 337

Ecoulement plastique des cristaux, mécanisme de dislocation 43
Effet acousto-électrique 225
— photo-électrique interne 219
— photovoltaïque 300
— piézo-électrique 229
— thermomagnétique 155
Efficacité du collecteur 314
— de l'émetteur 314
Einstein, relation d' 203
—, théorie de la chaleur massique du réseau cristallin 59
Electrons mis en commun 27, 76
Emission froide d'électrons 246
— stimulée de la lumière 307
— thermoélectronique 241
Energie d'échange 27, 170
— d'interaction des atomes 24
— de la liaison métallique 29
— d'un phonon 58, 61
— de répulsion suivant Born 24
— des vibrations normales 57
Enrichissement, région d' 329
Entraînement des électrons par des phonons 223
Epitaxie 270
Etat stationnaire 72
Ettingshausen, effet 154
Ewald, sphère d' 18
Exciton 214

Faraday, effet 190
Fente énergétique 140
Fermi, niveau de 91, 107, 112
—, quasi-impulsion de 193
—, surface de 92
Fermi-Dirac, fonction de 102
Ferrimagnétisme 173
Ferromagnétisme 159, 168
Film magnétique mince 183
Fluctuations 101
Fonction de densité des états 103
— — de répartition des modes 55
— d'onde 72
— de répartition 102
— — pour les fermions 102
— —, temps de relaxation de la 128
Fonction statistique complète 101
Force thermo-électromotrice 262
Franz-Keldych, effet 213
Frenkel, défauts de 32
Fréquence limite d'amplification en courant 323
— — d'une diode 291
Frontières des grains 45

Gain en courant d'un transistor 314
Gaz dégénéré de porteurs de charges 119
— non dégénéré de porteurs de charges 106
— de phonons 68
Générateur thermoélectrique 264
Gruneisen, formule de 67
Gunn, diode de 237
—, effet 237

Hall, constante de 149
—, effet 148 à 154
Hall-Schockley-Read, formule de 199
Hooke, loi de 42
Hystérésis magnétique 160, 179

Identité, période d' 14
Image électrique 239
— —, forme d' 239
Impureté 34
— acceptrice 97
— en position interstitielle 34
— — substitutionnelle 34
Indices de la direction 12
— de Miller 14
— de nœuds 12
— des plans 12
Injection de porteurs 256, 285
— — double 259
— monopolaire 260
Installation thermoélectrique frigorifique 267
Intégrale d'échange 28
Intégration, niveau d' 344
Interaction de dispersion 29
— d'échange 169
— des électrons avec les vibrations du réseau 222
— d'induction 31
— d'orientation 30
Inversion, région d' 329
Isolement diélectrique 347

Jonction *p-n* 269
— — à l'état d'équilibre 271
— —, méthodes de préparation d'une 269
Josephson, effet 146, 147

Kottrel, nuages de 40

Landau, niveau de 99
Landé, facteur de 164
Larmor, fréquence de 161
Laser 307
— à semiconducteurs 308
Liaison hétéropolaire 24
— ionique 24, 25
— métallique 29
— de Van der Waals 29
Lumière, amplification de la 307

Madelung, constante de 25
Maille élémentaire 8
Masque 346
Masse effective de l'électron 83
Matthiessen, règle de 137
Maxwell, fonction de 123
—, temps de relaxation de 207
Meissner-Ochsenfeld, effet 139
Mémoire à domaine magnétique 189
— à film magnétique 187
— à noyau en ferrite 181
Microélectronique 343
Microminiaturisation 343
Miller, indices de 14
Mobilité des porteurs de charges 129, 132
Mode 54
Montage à base commune 312
— à émetteur commun 318

Néel, température de 172
Niveau accepteur 97
— donneur 93, 95
— de Fermi 91, 107, 112
— d'impureté 93
— de Landau 99
— localisé 97
Nœud d'un réseau cristallin 8
Nombre de coordination 25
— d'états dans une bande d'énergie 85
— quantique magnétique 164
— — orbital 163
— — principal 73, 164
Non-dégénérescence, condition de 106

Ondes acoustiques de surface 230
— électriques dans le cristal 48, 222
— ultra-sonores, amplification d' 226
Optron 306

Paires de Cooper 140
— d'électrons 140
Paramagnétisme des diélectriques 163
Paramagnétisme du gaz électronique 165
Paramètres h, z et y des transistors 324
— d'une maille élémentaire 8
Paraprocessus 178
Pastille 350
Pauli, principe de 28, 85, 119, 146, 169
Peltier, coefficient de 266
—, effet 266
Perméabilité magnétique 158
Phonon 58
Photocellule 303
Photoconductivité 219
Photorésist 346
Photorésistance 219
Photothyristor 342
Phototransistor 325
Piège 196, 259
Planar, technologie 345
Poisson, équation de 206
Polarisation de l'onde 54
Population, inversion de 307
Porteurs de charges, génération de 191
— — en non-équilibre 191
Potentiel de déformation, constante de 224
Précession 161
Principe de Pauli 28, 85, 119, 146
Profondeur de pénétration du champ magnétique dans un supraconducteur 144

Quasi-impulsion de l'électron 82
— du phonon 70
Queue de la densité des états 116

Rapidité de fonctionnement d'une diode 284
Recombinaison bande à bande 194
—, coefficient de 194
— linéaire 195
—, lueur de 217
— de porteurs de charges 191, 195
— quadratique 195
Rectangularité du cycle d'hystérésis, facteur de 181
Redressemet, coefficient de 255
Remplissage des bandes en électrons 85
Rendement quantique de l'effet photo-électrique interne 219
Répartition des électrons suivant les vitesses 123
Réseau à base 11
— de Bravais 7
— cristallin 7

Réseau réciproque 18
Résistance différentielle négative 298
— diffusée 349
— à films 350
— théorique 41
Résistivité rémanente 137
Résonance cyclotronique 162
— magnétique électronique 166
— — nucléaire 167
— paramagnétique électronique 166
Réunion, coefficient de 302
Richardson, constante de 243
Richardson-Dushman, formule de 243
Rotation cohérente des spins 186
—, processus de 178

Saturation magnétique 159
Schéma équivalent d'une diode 290
— — d'un transistor 323
Schottky, barrière de 255
—, diode de 289
—, effet 245
Schrödinger, équation de 72
Semiconducteur amorphe 97
— compensé 117
— intrinsèque 88
— du type *n* 95
— — *p* 97
— vitrifié 342
Solution interstitielle 34
— substitutionnelle 34
Spectre de fréquences de vibrations du réseau cristallin 55
Spin de l'électron 163
Statistique, loi 100
Structure en mozaïque 44
Supraconducteur de deuxième espèce 144
Supraconductivité 138
—, nature de la 139
Susceptibilité magnétique 157
Syngonie 10

Technologie combinée 350
— hybride 350
— planar 345
Température caractéristique de Debye 61
— critique de supraconduction 138
Théorie de la chaleur massique du réseau cristallin, classique 58
— — — — — de Debye 60
— — — — — d'Einstein 59
— quantique de la conductibilité thermique 69
Thyristor 337, 342
Transfert des porteurs à travers la base, coefficient de 314
Transistor 312
— bipolaire 312, 347
— à dérive 322
— à effet de champ 327
— à jonction *p-n* de commande 334
— MOS à canal dopé 334
— unipolaire 333
Transition directe 211
— entre sous-bandes 215
— indirecte 212
Translation, période de 8
—, vecteur de 8
Travail de sortie 239
Trou 90
—, masse effective du 91
Tunnel, effet 76

Umklapprozesse 70

Vacance double 34
— triple 34
Van der Waals, liaison de 29
Vecteur de diffusion 18
— de réseau réciproque 18
Vibrations acoustiques 51
— normales des atomes 54
— optiques 51
— de point zéro 58
— thermiques des atomes 47
Vitesse de dérive 128
— de groupe de l'électron 80

TABLE DES MATIÈRES

Avant-propos 5

Chapitre premier. STRUCTURE DES CRISTAUX 7
§ 1.1. Réseau cristallin 7
§ 1.2. Désignation des nœuds, des directions et des plans dans le cristal. Indices de Miller 12
§ 1.3. Diffraction dans les cristaux. Réseau réciproque 15
§ 1.4. Méthodes expérimentales de détermination de la structure des cristaux 20
§ 1.5. Classification des cristaux suivant les types de liaison 24

Chapitre 2. IMPERFECTIONS DANS LES CRISTAUX 32
§ 2.1 Défauts ponctuels 32
§ 2.2. Défauts linéaires 36
§ 2.3. Mécanisme de dislocation de la déformation plastique des cristaux 41
§ 2.4. Défauts bidimensionnels 44

Chapitre 3. VIBRATIONS THERMIQUES DES ATOMES DE CRISTAL 47
§ 3.1. Vibrations et ondes dans un réseau unidimensionnel constitué d'atomes identiques 47
§ 3.2. Vibrations optiques 51
§ 3.3. Vibrations d'un réseau cristallin tridimensionnel 53
§ 3.4. Energie des vibrations normales. Phonons 57
§ 3.5. Chaleur massique des solides 58
§ 3.6. Dilatation thermique des solides 63
§ 3.7. Conductibilité thermique de réseau des solides 67

Chapitre 4. ÉLÉMENTS DE THÉORIE DES BANDES DES SOLIDES 72
§ 4.1. Niveaux d'énergie électroniques dans un atome isolé 72
§ 4.2. Mise en commun des électrons dans le cristal 74
§ 4.3. Fonction d'onde de l'électron dans le cristal. Zones de Brillouin 77
§ 4.4. Mouvement de translation des électrons dans le cristal 79
§ 4.5. Mouvement de l'électron dans le cristal sous l'action d'une force extérieure 81
§ 4.6. Remplissage des bandes en électrons et division des corps en métaux, diélectriques et semiconducteurs 85
§ 4.7. Semiconducteurs intrinsèques. Notions sur les trous 88
§ 4.8. Surfaces de Fermi des métaux 91
§ 4.9. Niveaux localisés dans la bande interdite 93
§ 4.10. Quantification de l'énergie de l'électron dans un champ magnétique. Niveaux de Landau 98

Chapitre 5. STATISTIQUE DES ÉLECTRONS ET DES TROUS DANS LES SEMICONDUCTEURS 100
§ 5.1. Fonction de répartition dans la statistique de Fermi-Dirac . . . 100
§ 5.2. Fonction de densité des états 103
§ 5.3. Concentrations d'électrons et de trous dans un semiconducteur 105
§ 5.4. Position du niveau de Fermi et concentration de porteurs de charges libres dans les semiconducteurs intrinsèques 107
§ 5.5. Statistique des électrons dans des semiconducteurs extrinsèques 110
§ 5.6. Statistique des électrons dans les métaux 118
§ 5.7. Chaleur massique du gaz électronique dans les métaux et les semiconducteurs 120

Chapitre 6. PHÉNOMÈNES CINÉTIQUES 123
§ 6.1. Répartition des électrons suivant leurs vitesses 123
§ 6.2. Dérive des porteurs de charges libres dans un champ électrique 127
§ 6.3. Mobilité des porteurs de charges libres. Variation de la mobilité avec la température 132
§ 6.4. Conductibilité électrique des semiconducteurs 135
§ 6.5. Conductibilité électrique des métaux 136
§ 6.6. Phénomène de supraconductivité 138
§ 6.7. Effets galvanomagnétiques 148
§ 6.8. Effets thermomagnétiques 155

Chapitre 7. PROPRIÉTÉS MAGNÉTIQUES DES SOLIDES 157
§ 7.1. Classification des solides suivant leurs propriétés magnétiques 157
§ 7.2. Nature du diamagnétisme. Résonance cyclotronique 161
§ 7.3. Nature du paramagnétisme 163
§ 7.4. Nature du ferromagnétisme 168
§ 7.5. Structure en domaines des ferromagnétiques 175
§ 7.6. Applications pratiques du ferromagnétisme 181

Chapitre 8. PORTEURS DE CHARGES HORS D'ÉQUILIBRE DANS LES SEMICONDUCTEURS 191
§ 8.1. Porteurs de charges en équilibre et hors d'équilibre. Quasi-niveaux de Fermi 191
§ 8.2. Durée de vie 193
§ 8.3. Equation de continuité 201
§ 8.4. Diffusion conjointe des porteurs hors d'équilibre 202

Chapitre 9. PHÉNOMÈNES OPTIQUES ET PLASMIQUES DANS LES SEMICONDUCTEURS 209
§ 9.1. Absorption de la lumière 209
§ 9.2. Emission d'un rayonnement lumineux par les semiconducteurs 217
§ 9.3. Photoconductivité des semiconducteurs 218
§ 9.4. Phénomènes acousto-électroniques 222
§ 9.5. Effet Gunn 231

Chapitre 10. CONTACT MÉTAL-SEMICONDUCTEUR 238
§ 10.1. Travail de sortie 238
§ 10.2. Emission thermoélectronique 241
§ 10.3. Différence de potentiel de contact 246
§ 10.4. Contact entre un métal et un semiconducteur 249
§ 10.5. Courants limités par la charge d'espace 256
§ 10.6. Force thermo-électromotrice 261
§ 10.7. Effet Peltier 266

Chapitre 11. DIODES À SEMICONDUCTEURS 269
§ 11.1. Etat d'équilibre d'une jonction *p-n* 269
§ 11.2. Capacité de barrière d'une jonction *p-n* 277
§ 11.3. Courants direct et inverse dans une jonction *p-n* 281
§ 11.4. Propriétés d'une jonction *p-n* lors du fonctionnement en impulsions et en hautes fréquences 284
§ 11.5. Claquage d'une jonction *p-n* 292
§ 11.6. Diodes tunnel 297
§ 11.7. Diodes à avalanche 298
§ 11.8. Phénomènes photo-électriques dans une jonction *p-n* 300
§ 11.9. Diodes luminescentes 305
§ 11.10. Lasers à semiconducteurs 307

Chapitre 12. PRINCIPES PHYSIQUES DE FONCTIONNEMENT DES TRANSISTORS 312
§ 12.1. Principe de fonctionnement d'un transistor bipolaire 312
§ 12.2. Paramètres et caractéristiques de sortie d'un transistor . . . 314
§ 12.3. Processus transitoires dans les transistors. Transistors à dérive 319
§ 12.4. Phototransistors 325
§ 12.5. Transistors à effet de champ 327
§ 12.6. Appareils à caractéristique courant-tension en S 336
§ 12.7. Dispositifs à semiconducteurs comme éléments des circuits intégrés 343

Annexes 352

Bibliographie 352

Index 353